Essays on Labour Economics in Developing Countries and the Role of the Public Sector

Ravi Somani

December 2018

A dissertation submitted in partial fulfilment

of the requirements for the degree of:

Doctor of Philosophy

University College London

Department of Economics

University College London

Declaration

I, **Ravi Somani**, confirm that the work presented in this thesis is my own.Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

Signature:

Abstract

This thesis presents a study of the internal functions of public-sector organisations in developing countries and how public-sector hiring policies interact with the wider economy, using Ethiopia as a case study.

The first chapter introduces the thesis.

The second chapter provides empirical evidence on how public-sector hiring policies impact education choices and private-sector productivity. I study the expansion of universities in Ethiopia and find that it leads to an increase in public-sector productivity but a decrease in private-sector productivity. The findings are rationalised by the following labour-market conditions, consistent with most developing country contexts: (i) entry into public employment is favoured towards those with a tertiary education; (ii) the public sector provides a large wage premium; (iii) entry into public employment requires costly search effort.

The third chapter uses novel survey instruments and administrative data to measure the individual-level information that bureaucrats have on local conditions and a field experiment to provide exogenous variation in information access. Civil servants make large errors and information acquisition is consistent with classical theoretical predictions. Bureaucrats have better information when: there are low costs of acquiring information; they are delegated authority over decision-making; there are organisational incentives to use operating information.

The fourth chapter analyses the introduction of a management-information system, which provides civil servants better access to information on their local environment, in the education sector. I find that the intervention significantly improves certain service delivery outcomes (enrolment), while other outcomes (schooling conditions) significantly worsen, consistent with the existence of multitasking in the bureaucracy.

The fifth chapter studies performance evaluations in the civil service. I use the dual-evaluation system in the Ethiopian civil service to measure supervisor bias and analyse its effects on service delivery, showing that supervisor bias is associated with efficiency losses in the public sector.

The final chapter concludes.

Impact Statement

Impacts Within Academia

The analysis presented in this thesis uses new data on public officials and public-sector organisations that I helped design and implement. Existing research on the functioning public sector has been limited by lack of detailed information on the characteristics of public-sector workers, their working environment, and productivity measures [Finan et al., 2017]. Research on the effects that public-sector hiring policies have on the wider economy and attempts at defining the optimal public-sector wage premium have also been limited by detailed individual-level data on public-sector workers. For the purpose of this thesis, I collected individual-level data on the characteristics and performance of a representative sample of public-sector workers in Ethiopia across five sectors of the service to help fill this gap.

Existing studies on the effects of expanding higher education have largely focused on developed country contexts, finding positive effects and leading to policy recommendations encouraging such expansions in developing countries [De Meulemeester and Rochat, 1995; Bloom et al., 2006]. The second chapter of this thesis ('Public-Sector Wages, Education Choices and Productivity') provides direct empirical evidence of the implications of expanding higher education in developing countries, taking advantage of new data on Ethiopian labour markets, including detailed information on the public sector, and a rapid expansion in public universities in Ethiopia. I provide evidence that such expansions can have negative effects for private-sector productivity under certain labour-market conditions, the first such evidence to my knowledge.

Impacts Outside Academia

Two chapters from this thesis ('Hierarch and Information' and 'Does Providing Information to Civil Servants Improve Service Delivery?') have featured in the 2018 Ibrahim Forum Report on Public Service in Africa, which highlights the latest research on public services in Africa.

The analysis from this thesis has also been part of a World Bank policy report titled 'Moving Further on Civil Service Reforms in Ethiopia', which has formed the foundation for future reforms in the civil service in Ethiopia, currently being drawn up.

The empirical evidence presented in the second chapter ('Public-Sector Wages, Education Choices and Productivity') highlighted a feature of labour markets in developing countries that was previously not documented – the fact that a large proportion of graduates in most developing countries are public employees. The significance of the

public sector for skilled labour markets in developing countries is important for designing education policy and public-sector hiring policies. This feature has now been incorporated in the Worldwide Bureaucracy Indicators published by the World Bank, as a result of this research.[1]

[1]https://datacatalog.worldbank.org/dataset/worldwide-bureaucracy-indicators [Accessed: 20th December 2018].

Acknowledgements

I would like to acknowledge my two main academic supervisors, Marcos Vera Hernandez and Antonio Cabrales, without whose guidance this research would not have been possible. Marcos has been the perfect primary adviser throughout my PhD – supporting and challenging my research ideas, providing invaluable advice over the progression of my research, and generally guiding me towards becoming a researcher in the field of economics. Antonio has been the ideal secondary advisor: he has been the source of inspiration of many ideas and has given me encouragement and confidence throughout the PhD. I feel extremely lucky to have had the support of both.

I am extremely grateful for my meetings with Fabien Postel-Vinay, who has helped me think in alternative ways about my research questions and has greatly supported my progression. I would also like to thank Imran Rasul for thoughtful meetings, helpful feedback, and critical advice over the course of the PhD.

My research focus and development has greatly benefitted from the support of Dan Rogger, who has involved me in fascinating projects, provided me with crucial advice towards their advancenment, and has greatly helped me to develop into an applied economist that researches The Bureaucracy.

In terms of funding, I am thankful to the Department of Economics at UCL for PhD scholarships. I am also thankful to the World Bank, the UK Department for International Development, and the International Growth Centre for project funding that allowed me to collect the data that made much of this research possible.

In terms of access to data and institutional support, I am indebted to the Government of Ethiopia and its representatives, who have been an integral part of my entire research agenda and professional development.

My research agenda grew from my experience as an economic advisor in the Ministry of Education, Ethiopia, through the Overseas Development Institute Fellowship. I am eternally grateful for this life-changing opportunity.

Contents

List of Tables

List of Figures

1 Introduction

1.1 The Importance of the Public Sector

State capacity is an important determinant of economic development [Besley and Persson, 2010; Acemoglu and Robinson, 2012*a*; Acemoglu et al., 2015] and recent literature has paid increasing attention to the effectiveness of government bureaucracies as a key feature of state capacity [Pepinsky et al., 2017; Finan et al., 2017].

The public sector is responsible for providing essential services to citizens, such as roads, infrastructure, public education, public health services, and public goods that are not efficiently provided by the private sector [Samuelson, 1954; Mas-Collell et al., 1995]. For this reason, the public sector accounts for a substantial proportion of total expenditure: the World Bank records that government expenditure accounted for 27% of global GDP in 2016 and the average across OECD countries was 41% in 2015 (the figure was 41% for the UK in 2017).[2] Researching the conditions under which the public sector is most efficient is therefore vital for economic efficiency and development.

The public sector also plays an important role in labour markets. Public employment accounts for 18% of employment across the OECD and around a quarter of paid employment across the sample of countries in the Worldwide Bureaucracy indicators published by the World Bank (the figure is 16% for the UK).[3] As I highlight in section 2, the public sector is a major employer of *skilled* individuals, especially in developing countries. For these reasons, the hiring policies of the public sector (the number of jobs, the wages attached to the jobs, and non-wage benefits) can have large impacts on the wider economy, as I discuss further in section 2.

1.2 Existing Literature on the Public Sector and Contributions

The question of how to improve the effectiveness of public organisations to transform scarce resources into vital public goods has been long-studied in the social sciences [Rose-Ackerman, 1986; Pepinsky et al., 2017]. The public sector generally operates under a different incentive environment relative to the private sector (in terms of asymmetric information, multiple tasks, multiple principals, and multiple levels of hierarchy) which affects the set of optimal contracts in a principal-agent framework [Dixit, 2002*a*].

[2]Figures sourced from https://data.worldbank.org/indicator/GC.XPN.TOTL.GD.ZS [Accessed: 20th December 2018]; OECD [2017]; and https://tradingeconomics.com/united-kingdom/government-spending-to-gdp [Accessed: 20th December 2018].

[3]See OECD [2017] and https://datacatalog.worldbank.org/dataset/worldwide-bureaucracy-indicators [Accessed: 20th December 2018].

There has been a growth in literature studying how changes in the incentive environment of bureaucrats can affect service delivery (see Finan et al. [2017] and Pepinsky et al. [2017] for reviews).

Regarding performance incentives, Dal Bó et al. [2013] provide experimental evidence on how changes in public-sector wages can determine the quantity and quality of applicants into the sector. Ashraf, Bandiera and Jack [2014] provides experimental evidence on financial and non-financial incentives for community-health workers and how these incentives interact with intrinsic motivation. Ashraf et al. [2016] provide experimental evidence on how varying the salience of career incentives relative to social incentives determine the types of applicants to the role and their subsequent performance in terms of household visits. Khan et al. [2015] experimentally vary incentives to tax collectors to analyse the effects on revenue collection and customer satisfaction. These studies find that performance incentives and incentives that highlight career advancement generally improve the quality of applicants for the job and their subsequent performance. On the other hand, Deserranno [2017] finds evidence that higher expected earnings discourage individuals with strong prosocial preferences from applying to positions in community health.

In terms of improving the monitoring of public officials, Callen et al. [2015, 2018] provide evidence of the effects of experimentally varying the extent of monitoring in the health sector in Pakistan. The authors find that the effect of monitoring significantly depends on the personality traits of the public employees and that the results of increased monitoring on bureaucrat performance depend on local politicians[4].

Rasul and Rogger [2016] and Rasul et al. [2018*a*] provide evidence of the importance of management practices for service delivery, highlighting the benefits of granting bureaucrats autonomy and the negative effects of performance incentives and targeting in the bureaucracy, explained partly by the existence of multitasking [Dixit, 2002*a*].

I extend this literature along several margins using the case of Ethiopia. Firstly, I provide empirical evidence on the effects of attracting talented individuals to the public sector on *private-sector* productivity. In section 2, I show that attracting the most educated and most intrinsically motivated individuals to the public sector generates productivity gains in the public sector but has detrimental effects on private-sector productivity as the most productive individuals select out. While there is an existing body of literature that studies the effects of public-sector contracts on the private-sector labour market [Quadrini and Trigari, 2007; Burdett, 2012; Albrecht et al., 2015; Gomes, 2015; Bradley et al., 2017; Gomes, 2017], I extend this literature firstly by incorporating detailed individual-level data on the characteristics and performance of

[4]The importance of the interaction between politicians and bureaucrats has also been highlighted by Iyer and Mani [2012], Rogger [2014], Nath [2015], and Pepinsky et al. [2017]

public-sector workers, which allows for better measures of public-sector productivity and effort. I also extend this literature by using plausibly exogenous variation in the supply of potential bureaucrats by exploiting a rapid expansion in higher education, since tertiary education is a pre-requisite for entry into the civil service. The expansion in public universities occurred over time and was geographically spread across Ethiopia, which in combination with the context of limited mobility across jurisdiction due to internal-migration policies, allows me to identify the effect of the expansion using narrow margins of variation. This identification strategy helps in overcoming the identification problem of endogenous selection into higher education and endogenous selection into public employment, which is a concern when using data that covers individuals only *after* they have selected into their occupation (and education level).

Secondly, to shed light on alternative interventions that can impact bureaucrat productivity, I analyse the effectiveness of providing civil servants better information on their operating environment in sections 3 and 4. The multidimensionality of tasks in the public sector and the difficulty of measuring individual-level productivity makes alternative interventions, other than performance incentives and monitoring, potentially attractive.

In section 3, I construct a novel way of measuring the information that individual civil servants have on their local jurisdiction and combine this with a field experiment that provides this exact information to a random subset of public organisations to analyse information acquisition in the public sector. The results show that civil servants, on average, have poor information on their operating environment. Analysing the heterogeneity in information shows that information acquisition is consistent with the classical theoretical framework on information acquisition, where civil servants have better information if they have authority over decision-making and if there are organisational incentives to gather and use operating information. The experiment reduces the marginal cost of information acquisition for a random sub-sample of organisations and this significantly improves information acquisition. However, heterogeneous treatment effects show that management practices (organisational incentives to gather and use operating information and organisational management practices that align the preferences of the principal and the agent) significantly mediate the effectiveness of the treatment.

In section 4, to link information acquisition directly to public-service delivery, I build a panel dataset which details the introduction of a management-information system in the education sector in Ethiopia. This intervention provides civil servants with better information on the local conditions in the jurisdiction. I exploit the phased rollout of the intervention, based on the district's distance to the regional capital and therefore accounted for by district-level fixed effects, for identification. I show that providing civil servants with better information on the current state of service provision

in their district allows them to re-optimise existing resource allocations (money, time, and implementation effort). This is consistent with the findings in section 3, that civil servants have poor information on current conditions in their jurisdiction, which makes it difficult to optimally allocate resources. The intervention leads to an increase in certain service delivery indicators (enrolment), but a deterioration in others (schooling conditions) in the short-run, highlighting the inherent nature of multitasking in the civil service, as bureaucrats distribute scarce resources across numerous tasks. I use detailed cost data to conduct a cost-benefit analysis of the programme, showing that the programme passes any reasonable cost-benefit test, but still requires substantial resources relative to existing levels of expenditure.

Thirdly, in section 5, I use the unique dual-reporting evaluation system of the Ethiopian bureaucracy to shed further light on performance evaluations in the bureaucracy. There is little evidence to date on the performance metrics and determinants of promotion within the bureaucracy [Rasul and Rogger, 2013; Goldfinch et al., 2013]. In this section, I use new individual-level data on the characteristics of civil servants, their performance evaluations, and the characteristics of the management and other organisational characteristics to provide further evidence on what determines progression in the civil service. Relating to the literature on subjective performance evaluations, I find evidence of supervisor bias in the evaluations and find that this bias is associated with efficiency losses in terms of worse service delivery [MacLeod and Malcomson, 1989; Prendergast and Topel, 1993; Baker et al., 1994; Prendergast and Topel, 1996; Prendergast, 2002; Frederiksen et al., 2017]. I use the dual-reporting system to produce an individual-level measure of supervisor bias and show that this measure: (i) changes with the gender of the employee and the manager (even though the performance of the employee is not significantly associated with gender); (ii) feeds into the wages of civil servants, and significantly more so when managers have greater autonomy; (iii) is correlated with alternative measures of bias (for example, using wage residuals after controlling for observable characteristics). I then create an organisational-level measure of supervisor bias and show that this is significantly negatively correlated with service delivery outcomes, even after conditioning for employee characteristics, manager characteristics, management practices, jurisdiction characteristics, and sector characteristics. I show that the effect of supervisor bias on service delivery is even more negative in environments where preferences are diverse or where there is worse monitoring – where the productivity of an individual civil servant is harder to measure and hence the probability of the supervisor being punished for biasing the reports lower.

In section 6, I provide concluding comments and suggestions for further research.

2 Public-Sector Wages, Education Choices and Productivity

Skilled labour markets in developing countries are characterised by public employment. The most common occupation for graduates in developing countries is in the civil service: around *half* of tertiary-educated employees work in the public sector in the poorest countries.[5] As such, public-sector wages and hiring policies determine the expected returns to higher education and education choices, as well as occupation choices, with consequences for productivity in both the public sector and the private sector. This paper studies how public-sector hiring policies interact with expansions in higher education and the implications for education decisions, labour-market outcomes and productivity.

I provide micro-level evidence on the impacts of expanding higher education on education decisions, labour-market outcomes and productivity in Ethiopia, where 58% of all post-secondary-educated individuals are employed in the public sector. I exploit the rapid expansion of public universities in Ethiopia during the 2000s using a comprehensive new dataset that I constructed. The dataset includes detailed information on public-sector workers' characteristics and performance from a survey that I helped design and implement, as well as representative data on education attainment and labour-market outcomes. I find evidence that attractive and rigid public-sector wages for graduates lead to the selection of the most productive individuals into higher education and public-sector job search, generating gains in public-sector productivity, while private-sector productivity decreases.

As I show in a simple model that restricts access to public employment to graduates only, the expansion in public universities leads to the selection of the most productive individuals into higher education and (costly) public-sector job search if: (i) there is positive correlation between public-service motivation and ability; and (ii) the public-sector wage premium is large.[6] The context of the expansion in Ethiopia, combined with the detailed information that I have on the characteristics of public-sector workers, allows me to explore both of these dimensions in the empirical analysis. The expansion in public universities in Ethiopia was phased in over time and was geographically spread for political considerations regarding regional equity [Ashcroft and Rayner, 2011]. The interaction of time (cohort) and geographical variation in the rollout of public universities allows me to categorise individuals into groups based on their exposure to the

[5]Based on the Worldwide Bureaucracy Indicators, World Bank. Estimates from the Living Standards Measurement Surveys of the World Bank show that 53% of *all* post-secondary-educated individuals work in the public sector in Tanzania, 40% in Nigeria 36% in Bulgaria, and 20% in Uganda. Girsberger and Meango [2017] shows that 40% of tertiary-educated individuals across 6 West African capital cities work in the public sector.

[6]Public-service motivation reduces the cost (disutility) of public-sector job search.

establishment of new public universities, in the spirit of Card and Lemieux [2001], Duflo [2001], and Atkin [2016]. The geographical variation in the establishment of universities, combined with the limited mobility of individuals across districts (labour markets) due to rural-development policies in Ethiopia, allows me to estimate the effect of the expansion across different labour markets that are heterogeneous in the public-sector wage premium.

Firstly, I find that the expansion in public universities had a positive effect on education attainment, leading to a 3.4 percentage-point increase in the proportion of individuals with a tertiary education (31% of the counterfactual mean). However, the take-up of tertiary education is significantly mediated by the public-sector wage premium in the district. There is no effect of the expansion on education attainment in districts with the lowest public-sector wage premium and effect sizes of a 15-percentage-point increase in the districts with the largest public-sector wage premium. These results demonstrate the role of public-sector contracting policies in determining education decisions and human-capital accumulation in the economy.

Secondly, I find that the university expansion leads to a decrease in the employment rate in the local economy, driven by a decrease in the employment rate of tertiary-educated individuals (6 percentage-points). Individuals exit the formal private sector to enter higher education and public-sector job search; and do so more in areas with a higher public-sector wage premium. This result is consistent with individuals 'queueing' for attractive public-sector jobs rather than accepting private-sector employment, which generates unemployment in the local labour market [Quadrini and Trigari, 2007; Albrecht et al., 2015; Gomes, 2015]. This finding is rationalised by costly search for public employment, which prevents individuals from working while searching for public employment [Abebe et al., 2016; Chamberlain, 2015, 2017].

Thirdly, I provide evidence that private-sector wages decrease significantly as a result of the expansion in higher education (0.13 log-points). This is consistent with the most productive individuals leaving the private sector to enter higher education and search for public employment. I develop a conservative estimate of the decrease in private-sector productivity using wage residuals which suggests that productivity in the private sector falls by 2 percentage-points as a result of the expansion in higher education. Estimates from quantile regressions across the wage distribution show that the negative effects on private-sector wages are concentrated at the top end of the wage distribution, providing further evidence that the most productive individuals exited the private sector after the expansion. Since public-sector wages do not respond to the expansion in public universities, the public-sector wage premium increases further. I show that the increase in the public-sector wage premium is largest in areas with already high premia, exacerbating existing inequalities between private-sector and public-sector

earnings. The increase in the public-sector wage premium increases the returns to higher education, as the majority of individuals with a higher education work in the public sector and the majority of less-skilled individuals work in the private sector.

Finally, I estimate the effects of the expansion on public-sector productivity. Since measuring public-sector productivity from wages is problematic due to their rule-based nature, I use purposefully-collected data on the characteristics and performance of public-sector workers to measure the impacts of the expansion on the public sector. I find evidence that public-sector productivity increases as a result of the expansion in higher education. Bureaucrats that enter public service after the expansion display significantly higher levels of public-service motivation (0.7 standard deviations) and mission alignment (38.4 percentage-points). Furthermore, they work significantly more hours per week (6 hours), for the same wage, and earn higher scores in their performance evaluations (3.76 percentage points). Yet, I find that only public-sector organisations with strong management practices related to screening are able to recruit more motivated bureaucrats after the expansion, highlighting the importance of screening technologies for identifying talent among a growing pool of graduates.

On the whole, I find that that the expansion in higher education in Ethiopia led to a reduction in private-sector productivity and an increase in public-sector productivity. I find that the effects of the expansion in higher education on education attainment are significantly mediated by local labour-market conditions, namely the public-sector wage premium. This finding highlights the complementarity between higher-education policies and public-sector hiring policies. Relatively attractive and fixed public-sector wages for the tertiary educated ensure that expansions in higher education draw the most productive individuals out of the private sector and into higher education and public-sector job search. Since public-sector hiring policies determine the education choices of individuals and the distribution of skills in the economy, they have the potential to impact the future growth potential of the economy.

A natural concern regarding identification is that the expansion in public universities is endogenous to local labour market conditions. I use the 1999 Labor Force Survey to show that the districts that receive a public-sector university are balanced, relative to those that do not, prior to the expansion. Another concern is that district-specific changes across cohorts (e.g. mean reversion) are driving the results. I run a control experiment similar to Duflo [2001] to test whether there is evidence of this in the data and find no such evidence. A further concern is that the establishment of a new university brings about demographic changes across cohorts and districts, for example, through migration. I run a series of placebo regressions using the 2013 Labor Force Survey and the Ethiopian Civil Servants Survey 2016 to show that the results are not driven by such changes. Finally, I turn to additional data sources, using the World

Values Survey and the Living Standard Measurement Survey, to rule out alternative mechanisms driving the results, such as financial constraints, changes in the composition of the public-service motivation of graduates, and alternative sources of income (estimating the effects on consumption).

This paper contributes to the literature that studies the impact of public-sector contracting policies on the wider labour market, with new micro-level evidence [Hörner et al., 2007; Burdett, 2012]. While some of the unintended effects of attractive and exogenously determined public-sector wages are documented, such as unemployment, crowding-out of private-sector job-creation, and business-cycle volatility [Quadrini and Trigari, 2007; Gomes, 2015; Albrecht et al., 2015; Bradley et al., 2017], I provide evidence of an additional consequence: attractive public-sector wages reserved for graduates determine education choices and the distribution of skills in the economy. This determines the stock of human capital in the economy with short-run consequences for private-sector productivity as well as long-run consequences for economic growth. Further, I show that attractive public-sector wages for graduates incentivise high-skilled labour to queue for public-sector employment, providing an alternative explanation for the phenomenon of high unemployment among the tertiary-educated in developing countries [Todaro, 1969; Fan and Stark, 2007; Girsberger and Meango, 2017].

I also contribute to the literature on the importance of education for economic development [Mankiw et al., 1992; Barro, 2001]. While the implications of expanding general education have been highlighted in the existing literature (for example, Duflo [2001] and Duflo [2004]), the evidence on expanding tertiary education is much thinner [Glewwe, 2002; Li et al., 2014; Oppedisano, 2014; Boccanfuso et al., 2015]. This is despite the fact that globally more than one-fifth of public expenditure on education is allocated to tertiary education and that more is spent per pupil in tertiary education than primary or secondary [Glewwe and Kremer, 2006]. Furthermore, I provide evidence that reducing the cost of access to schooling in higher education also increases education attainment, in line with much of the existing literature on general education [Duflo, 2001; Card, 2001; Glewwe and Kremer, 2006; Filmer, 2007].

Finally, I contribute to the literature on the productivity of public officials [Finan et al., 2017]. While the positive effects of incentives, monitoring and management have been shown to increase the productivity of public officials [Dal Bó et al., 2013; Ashraf, Bandiera and Jack, 2014; Callen et al., 2015; Ashraf et al., 2016; Rasul and Rogger, 2016], I show that screening based on education, combined with attractive public-sector wages, can lead to improvements in public-sector productivity. I show that this can come at a cost to private-sector productivity, however, as the most productive individuals select into public-sector job search. I extend this line of literature in two ways: firstly, I capture direct measures of the performance of public officials across a

range of service-delivery sectors; secondly, I connect the productivity of public officials to the wider economy and document the unintended consequences of attracting talent to the public sector for private-sector productivity.

The next section describes the dataset. Section 5.2 outlines the context, the expansion in public universities, and the definitions of exposure. Section 4.3 presents the conceptual framework. Section 4.4 discusses the identification strategy and the validation checks. Section 2.5 presents the results from the empirical analyses. Section 2.6 explores alternative mechanisms that could be driving the results. Section 2.7 concludes.

2.1 Data Sources

2.1.1 University Database

I extracted data on university establishments and enrolment from the Ministry of Education Education Statistics Annual Abstracts from 1999/2000 to 2013/14 academic years. I then geo-located each of the public universities using the Ministry of Education's website and the website of each of the public universities individually, linking universities to district administrations.[7] I also created an exhaustive list of private universities from both the Ministry of Education Education Statistics Annual Abstracts and the Ministry of Education website.[8] For all of those private universities outside of Addis Ababa, I matched the location of the private universities to district administrations, including for secondary campuses of private universities.[9] Since Addis Ababa, the capital city, has had a university since 1950, it is difficult to explore the effect of its establishment on the labour market with recent data (individuals aged 16 at the time of the establishment of Addis Ababa University are aged 79 in the Labor Force Survey 2013). I therefore exclude Addis Ababa from the analysis. I also exclude the Ethiopian Civil Service College, the Defense University College, and Kotebe Teachers Education College, due to their special nature, since individuals entering these particular institutes are different from those entering regular undergraduate degree programmes.

2.1.2 Ethiopian Labor Force Survey (LFS) 2013

The LFS 2013 data is the primary dataset for the analysis on education attainment, employment, and private-sector wages and productivity. The LFS contains district codes that allow me to link individuals to district administrations. I merge the data from the LFS with the data from the Ministry of Education. The LFS 2013 data

[7] http://info.moe.gov.et/pubuni.shtml [Accessed: 16 July, 2018].
[8] http://info.moe.gov.et/colls.shtml [Accessed: 16 July, 2018].
[9] The majority of private-university enrolment is in Addis Ababa (92% in 2013/14).

contains information on employment, wages, occupation, sector (public/private), and demographic characteristics of respondents for a representative sample of individuals in the economy. I restrict the data to include only those aged between 21 and 65, i.e. of working age and at least the minimum usual age to acquire a tertiary education.

2.1.3 Ethiopian Labor Force Survey 1999

I use the 1999 Ethiopian Labor Force Survey to compare, prior to the rapid expansion in public universities during the 2000s, districts that end up with a public university and districts that do not. The data contains information on employment, occupation, sector (public/private), and demographic characteristics of respondents for a representative sample of individuals in the economy (not wages).

2.1.4 Ethiopian Civil Servants Survey 2016 (ECSS)

The ECSS was conducted between June 2016 and September 2016 and I was involved in its design and implementation, working in collaboration with the Government of Ethiopia and the World Bank. The data contains information on a representative sample of civil servants across Ethiopia. The survey captures the basic characteristics of bureaucrats (age, tenure, gender, education), the management practices under which they operate [Bloom and Van Reenen, 2007*a*; Rasul and Rogger, 2016], the public-service motivation of employees [Perry, 1996], the perceptions of employees regarding the benefits of working in the public sector (questions on mission alignment; questions on motivation), the time-use of civil servants (hours worked in a typical week), and performance-evaluation data for a subset of civil servants [Somani, 2017]. Since this paper is focused on the effects of new universities on local labour markets and bureaucrat selection, I restrict the analysis to district-level civil servants, as district-level organisations are geographically spread and the first entry point into the service for the majority of bureaucrats. For consistency with the LFS, I restrict the sample to only those aged between 21 and 65.

2.1.5 Ethiopian Census 2007

The census of 2007 is used to capture basic demographic characteristics at the district level.

2.1.6 World Values Survey (WVS) 2007

To investigate alternative mechanisms, I use wave 5 of the World Values Survey (WVS), to investigate changes in public-service motivation for all individuals, including those

not working in the public sector. I use Wave 5 due to the inclusion of Ethiopia for this wave.[10] For consistency, I restrict the sample to only those aged between 21 and 65.

2.1.7 The Living Standards Measurement Survey (LSMS)

To explore whether financial constraints play a role or whether individuals experience changes in consumption (not just wages), I use the Living Standards Measurement Survey conducted by the World Bank. I use Wave 3, the latest round, to gather information on measures of wealth (parent education and household assets) and expenditure on consumption in the last month.[11] For consistency, I restrict the sample to only those aged between 21 and 65.

2.2 Context

2.2.1 Institutional Background

Ethiopia is Africa's second-most populous country, home to 100 million people or 10% of the population of Sub-Saharan Africa. The GDP per capita is 707 USD and Ethiopia ranks 173 out of 189 on the Human Development Index. Its ethnically diverse population is dispersed across over a million square kilometres, such that population characteristics vary widely across space. In addition to geographic and linguistic constraints, mobility across districts is constrained by policies that limit the transfer of land rights and restricted access to public services for migrants. Like other developing countries, government expenditures represent a significant fraction of GDP (18%) and the public sector is a large employer of formal workers. Corruption is less prevalent than in most of sub-Saharan Africa, but government effectiveness is seen as relatively weak overall. Never fully colonised, Ethiopia's government is organised as a developmental state, with a relatively vertical and rigid hierarchy.

2.2.2 The Education System in Ethiopia

The structure of the education system in Ethiopia is as follows: primary level consists of grades 1-8, where the 'normal' age of students is 7-14; general secondary level consists of grades 9-10, after which there is a national examination in which students need to pass a minimum of 5 subjects to proceed to higher secondary education; higher secondary level consists of grades 11-12, after which there is a national university entrance examination, which determines whether a student is eligible to enter university. An alternative route

[10] The data and documentation can be found at:
http://www.worldvaluessurvey.org/WVSDocumentationWV5.jsp [Accessed: 16 July, 2018].

[11] Data and documentation can be found: http://microdata.worldbank.org/index.php/catalog/2783 [Accessed: 16 July, 2018].

after grade 10, rather than higher secondary level, is to enter technical or vocational education, focused on specific career tracks. To give some sense of the opportunity cost of study at secondary and higher education level, the average monthly salary for those aged 15-18 without a secondary level education is 13 USD per month and for those aged 19-22 without a post-secondary level education is 23 USD per month.

To enter higher education, students need to achieve a minimum score in the national exam at the end of grade 12, which is decided centrally, based on the number of university places available. As the number of public universities expands, therefore, the threshold score required for entry decreases, allowing a greater proportion of students to enter. For example, the percentage of students achieving a score over 350, the threshold for entry into higher education, increased from 30% in 2010 to 45% in 2014, 50% in 2015. During the application process for entry into higher education, students put down their preferences for the universities that they wish to attend and the broad field that they wish to study. The Ministry of Education applies an allocation mechanism that gives priority to students' first choices in university and field of study. If courses and fields of study are oversubscribed, the Ministry of Education allocates students to universities and courses based on their second choices and so on.

The Labor Force Survey 2013 data shows that 19% of the population aged 21-65 have completed primary education, 8% have completed secondary education, and 6% have a post-secondary education.

2.2.3 The Expansion in Higher Education in Ethiopia

To bring about the skills in the population required to achieve the national development goals, the Ethiopian government heavily invested in the higher education system in recent years, attaching a disproportionately high share of the education budget to higher education [Ravishankar et al., 2010]. For example, 45.9% of the education budget was allocated to higher education in 2014/15, the latest year of budget data available, around 2% of GDP.

The Ethiopian higher-education sector witnessed a rapid expansion during the 2000s, as the number of public universities increased from 8 to 33 between 1999 and 2013, as shown in figure 1 (left) and table 6, with further expansions planned. The expansion can also be seen in the growth in enrolment in figure 1 (right). The locations of the new universities were chosen to ensure regional equity, in terms of universities per region, and based on distance to existing universities, such that no two universities were 'too close' together [Ashcroft and Rayner, 2011]. Therefore, the expansion in universities incorporates both time variation (table 6 and figure 1) and spatial variation (figure 2). Almost all students attend a public university within their home district (91% of

post-secondary educated individuals).

To align the curriculum with the national development strategy to achieve middle-income status by 2025, from 2008 the Ethiopian government required that 70% of all students should study subjects related to science and technology [Federal Democratic Republic of Ethiopia, 2010, 2016; Ashcroft and Rayner, 2011]. As can be seen from figure 13, this policy saw that the growth in engineering and technology graduates outstripped that of other areas of study.

In terms of teaching quality indicators for the 2013/14 academic year, I find no evidence that the newer universities have a lower proportion of teachers with at least a bachelor's degree and I find no evidence that the new universities have worse pupil-teacher ratios, suggesting that there is no observable evidence on these dimensions that the newer universities provide lower levels of teaching quality.

2.2.4 Labour Markets in Ethiopia: High-Skilled and Low-Skilled

2.2.4.1 The Importance of Public Employment for the Skilled in Developing Countries

The high-skilled sector (those with a post-secondary education) accounts for 6% of the working age population in the LFS 2013 sample or 11% of formal employment. The public sector is the major employer of the high-skilled in Ethiopia, representing 58% of all post-secondary-educated individuals (figure 3, left), or 68.9% of all *employed* post-secondary-educated individuals in the LFS sample, or 80% of all *formally employed* post-secondary-educated individuals.

The 'public sector' here is defined as all employees of federal, regional, district, and local level government agencies and state-owned enterprises, including 'frontline' civil servants, such as teachers and doctors. In terms of industries of work, individuals working in the public sector mostly work in public administration (57%), finance (16%), and agriculture (9%). 90% of public officials in Ethiopia work at the non-federal (decentralised) level.

The importance of the public sector for high-skilled labour in Ethiopia is not dissimilar to other developing countries: estimates from the Living Standards Measurement Surveys of the World Bank show that 36% of *all* post-secondary-educated individuals work in the public sector in Bulgaria, 40% in Nigeria, 53% in Tanzania and 20% in Uganda. Girsberger and Meango [2017] shows that 40% of tertiary-educated individuals across 6 West African capital cities work in the public sector. Figure 14, using data from the Worldwide Bureaucracy Indicators, shows the proportion of employed tertiary-educated individuals working in the public sector and shows that this figure tends to be larger for poorer countries, with an average of around 50% in the lowest-

income countries, in Sub-Saharan Africa. Figure 15 shows figure 8 from Finan et al. [2017], which shows the secondary and tertiary education premium in the public sector by the GDP per capita of the country. The figure shows that the public sector consistently demands higher levels of education than the private sector (all observations in the sample have a positive premium) and that this premium increases as GDP per capita decreases.

A complementary feature of the high-skilled labour market is the low tertiary-education wage premium in the private sector relative to the public sector, shown in figure 3 (right). The figure shows that the log hourly wage tertiary education premium, controlling for all functions of age, gender, and district fixed effects, is 0.3 log points in the private sector but 0.7 log points in the public sector. This suggests that there are relatively few attractive jobs in the private sector for skilled individuals. For those with a post-secondary education working in the private sector, the main activities are agriculture (37%), manufacturing and construction (18%), commerce (14%), education or health services (9%), and community services (9%).

2.2.4.2 Low-Skilled Labour and the Private Sector

As can be seen in figure 3 (left), the public sector is much less important in terms of proportional employment for individuals without a tertiary education ('low-skilled'): 2% of the non-post-secondary educated work in the public sector, or 2.6% of employed low-skilled labour, or 6.5% of formally employed low-skilled labour. For individuals with very low skills (primary education or less), 1% work in the public sector. For those with a primary education or less working in the private sector, agriculture is the main line of work for the very low skilled (79%), followed by manufacturing and construction (7%), community services (6%) and commerce (6%). For individuals with intermediate skills (secondary level education), 17% work in the public sector. For the secondary educated in the private sector, agriculture is still the main activity for these individuals (47%), followed by commerce (18%), manufacturing and construction (17%), and transport and communication (9%).

As is documented in Sutton and Kellow [2014], private-sector enterprise in Ethiopia has a small number of large firms (43 firms with more than 500 employees and 408 firms with 50-50 employees), highlighting the prevalence of small firms in the private sector. The Enterprise Survey data from the World Bank shows that the most-cited serious obstacle to firms is access to finance, selected by more than 40% of firms.[12] This lack of growth in private-sector firms helps explain low wages and the low returns to education in the private sector.

[12] http://www.enterprisesurveys.org/data/exploreeconomies/2015/ethiopia#firm-characteristics [Accessed: 24th October, 2018]

2.2.5 Migration and Mobility

Policies that restrict the transfer of land rights and that attach access to public services to strict district residence conditions limit migration in Ethiopia [Dorosh and Schmidt, 2010]. Part of the reasoning for these policies is specifically to ensure that economic development is not concentrated regional, but rather takes place broadly across the country [Federal Democratic Republic of Ethiopia, 2003; Oswald and Schenker, 2010].

The first such policy is related to land policy. Land is officially owned by the government and leased to individuals to farm. Local governments manage land allocation rights and these are linked to active farming of the land, providing adequate care to the land and remaining a resident of the local neighbourhood [Dorosh and Schmidt, 2010]. The second policy is related to access to public services. A migrant needs to live in the new district for 6 months before being able to register their address and receiving an identification card that recognises the new address. The identification card is required to access public services, hence migrants must forego access to public services for 6 months while they relocate. Even after 6 months of residence, a migrant may still be unable to apply for an identification card, since applications are restricted to individuals that own a property in the district. An alternative route is for the residence to migrate through social networks and family ties that own property in the destination district and can apply for an identification card on behalf of the applicant [Dorosh and Schmidt, 2010].

International migration out of Ethiopia is very low (1.2% of the population). Internal migration within Ethiopia is typically quite low with 4% of all individuals having migrated for work and less than 1% for education.[13] Internal migration rates are typically higher for those with a post-secondary education: 4.6% have migrated for work (this figure excludes public-sector transfers) and 9% have migrated for education .

The limited mobility of individuals across districts and the emphasis on districts as development centres ensures that districts are largely isolated labour markets.

2.2.6 Public-Sector Hiring

Public-sector hiring as a proportion of the working-age population is relatively constant over time, with around 5% of the population or 6% of employed individuals working in the public sector. This is relatively constant over different cohorts as shown in figure 16. The fact that public-sector hiring is constant as a proportion of the population over time shows that there are no sudden changes in labour demand in the public sector

[13]This figure is higher when one includes forced migration. I exclude this from the figures as this is not an endogenous choice or part of endogenous selection by the migrant, but rather due to natural disasters, conflict, or arranged marriage.

between cohorts, and this is true when one looks at districts with public universities and districts without public universities separately.

2.2.6.1 Costly Public-Sector Job Search

An undergraduate degree is considered a pre-requisite for entry into the professional civil service and 82% of professional civil servants currently have a tertiary education. Most civil servants heard about their job through advertising boards (45%), informal invites or transfers (39%) and newspaper adverts (13%). Job search is a costly process in Ethiopia and public-sector hiring is typically a slower process than private-sector recruitment [Chamberlain, 2015, 2017; Abebe et al., 2016]. Travelling to advertising boards in the main town repeatedly, purchasing (or renting) newspapers, travelling to offices to submit CVs and applications, and attending interviews are costly (and uncertain) processes for individuals interested in public-sector jobs. As such, undertaking this process while holding a full-time or formal private-sector position is difficult.

2.2.6.2 Public-Sector Wages

The public-sector wage in Ethiopia is determined by a grading system within the civil service, which is a function of education, tenure, and achieving a minimum level of performance. The wages are a direct mapping from the grade of the civil servant and are reviewed centrally by the Federal Ministry of Public Service and Human Resource Development and any changes must be ratified by the Council of Ministers, hence the wage-determination function is the same for all district governments [Markos, 2013; World Bank, Forthcoming]. The wage is relatively constant over time, having only been adjusted five times between 1991 and 2014 [Addis Fortune, 2014].

2.3 Theoretical Framework

In this section, I develop a model of selection into higher education and public-sector job search to analyse how the public-sector wage premium and establishment of a public university will affect the characteristics of entrants into higher education and public-sector job search.

Individuals are infinitely-lived, risk-neutral, and heterogeneous in two dimensions: ability (α) and public-service motivation (ψ). α is distributed according to the cumulative distribution function $F(\alpha)$ and density $f(\alpha)$. I present two cases: in the first case α and ψ are independently distributed among the population; in the second case α and ψ are correlated and, to simplify matters, following Dal Bó et al. [2013], I consider the extreme case where all types (α, ψ) are contained in the graph of function $\psi = m(\alpha)$.

This correlation is not supposed to provide general insights but to provide contrasts to the independent case to help guide and interpret the empirical analysis.

Private-sector firms are homogeneous and pay wages equal to the marginal productivity of labour, which is a linear function of ability in this case: $w = \pi\alpha$.[14] Private-sector firms observe the ability of the individual. I acknowledge that this version of the private-sector market is highly simplified and that a structural estimation of the model would benefit from a more sophisticated labour-demand function and the existence of market frictions; I impose the simplifications to offer a contrast to the exogenous wage-posting of the public sector and to help interpret the empirical analysis.

Public-sector organisations pay exogenous wages w^G and screen individuals *with a higher education* at the recruitment stage. The public-sector job-finding process is costly, hence only individuals with sufficiently high public-service motivation are willing to endure the process [Chamberlain, 2015, 2017]. An individual is offered a public-sector job based on her public-service motivation (ψ) and some luck (ϵ): $P(\psi + \epsilon \geq \underline{t})$, where $\underline{t}$ is an exogenous threshold and where ϵ is unbounded and distributed according to cumulative distribution function $G(\epsilon)$ and density $g(\epsilon)$.[15] All job offers are accepted. In this case, if the individual doesn't receive a public-sector job offer, she works in the private sector. I introduce unemployment in the following sub-section.

Individuals pay a one-off cost of access (C) to enter higher education, after which they may apply for public-sector employment. The cost of access to higher education is unrelated to ability but I consider the case where the cost of access decreases with ability as an extension. Let W^P represent the value of private-sector employment and W^G represent the value of public-sector employment. I assume no separations or transitions after the education choice is made. Time is continuous and discounted at rate r.

The set of individuals that enter higher education have values of α and ψ such that:

$$rW^P \leq P(\psi + \epsilon \geq \underline{t})rW^G + [1 - P(\psi + \epsilon \geq \underline{t})]rW^P - C \tag{1}$$

Which simplifies to:

$$\pi\alpha \leq [1 - G(\underline{t} - \psi)]w^G + G(\underline{t} - \psi)\pi\alpha - C \tag{2}$$

[14] The firm maximises profits from labour, $\pi\alpha L - wL$, with respect to L.

[15] For example, her public-sector job-search effort increases in public-service motivation. There are also efficiency-based arguments for screening on public-service motivation [Besley and Ghatak, 2005]. For example, public-service motivation is significantly positively associated with hours worked in the sample of civil servants in the Ethiopian Civil Servants Survey data, even after controlling for gender, education, tenure, and civil service seniority (grade). It is assumed that private-sector organisations can observe ability since there are typically better measures of productivity in the private sector, meaning that the private firm can learn ability over time, while this is not the case in the public sector [Dixit, 2002*a*].

Proposition 1: The most able individuals select into higher education if there is positive correlation between ability and public-service motivation under the following conditions:

(a) *If the public-sector wage, w^G, is set sufficiently high relative to the private wage*

(b) *Private-sector firms demonstrate low productivity so that π is 'small'.*[16]

Proof:

In the case where α and ψ are independently distributed in the population, there is only negative selection, in terms of ability, into higher education. At the same time, there is positive selection, in terms of public-service motivation, into higher education. See Theory Appendix for proof.

In the case where all types (α, ψ) are contained in the graph of function $\psi = m(\alpha)$, taking functions of α to one side, inequality 2 becomes:

$$\pi\alpha + [w^G - \pi\alpha][G(\underline{t} - m(\alpha))] \leq w^G - C \tag{3}$$

Let $k(\alpha) := \pi\alpha + [w^G - \pi\alpha][G(\underline{t} - m(\alpha))]$, where $k'_\alpha(\alpha) = \pi[1 - G(\underline{t} - m(\alpha))] - m'(\alpha)g(\underline{t} - m(\alpha))[w^G - \pi\alpha]$.

The most able individuals select into higher education and public-sector job search if $k'_\alpha(\alpha) < 0$, which is the case if and only if $m'(\alpha) > 0$ – when there is positive correlation between ability and pubic-service motivation – and the public sector pays higher wages than the private sector so that $[w^G - \pi\alpha] > 0$. This is illustrated in figure 4, which plots $k(\alpha)$ and $w^G - C$ under the case when $k'_\alpha(\alpha) < 0$ (left) and $k'_\alpha(\alpha) > 0$ (right), where the decision to enter higher education is based on inequality 3 and $k(\alpha)$ is the left-hand side of the inequality.

If $m'(\alpha) < 0$, then $k'(\alpha) > 0$, leading to negative selection into higher education in terms of ability (figure 4, right). If $m'(\alpha) > 0$, then the sign of $k'(\alpha)$ becomes ambiguous. $k'(\alpha)$ is negative if $\pi[1 - G(\underline{t} - m(\alpha))] < m'(\alpha)g(\underline{t} - m(\alpha))[w^G - \pi\alpha]$. This is the case if $m'(\alpha) > 0$, if π is close to zero and if $[w^G - \pi\alpha]$ is large. Under these conditions, the most able individuals select into higher education and public-sector job search (figure 4, left).

To highlight the importance of the public-sector wage premium in education decisions, suppose that the public-sector wage premium is zero, so that $w^G = \pi\alpha$. From

[16]This reflects the context described in section 5.2: private-sector wages are typically low, private-sector firms operate in low-productivity industries (agriculture and low-skilled services), and the higher-education wage premium in the private sector is significantly lower than the public sector, suggesting less value placed on skills.

inequality 2 it can be seen that nobody will attend higher education: the potential benefit of attending higher education is now zero but there is still a cost of attending.

Suppose that the private wage is very low so that the public-sector premium is high. For simplicity, suppose that $\pi = 0$. Then, since $w^G > 0$, we are in the case where the sign of $k'_\alpha(\alpha)$ depends entirely the sign of $m'(\alpha)$. If $m'(\alpha)$ is positive, then the most able individuals select into higher education. Hence, when the public-sector premium is high and when public-service motivation is positively correlated with ability, the most able individuals will select into higher education.

Extending the framework to allow the cost of access to higher education to be a decreasing deterministic function of ability, for example because of the lower psychological cost of undertaking a degree or increased access to scholarships for the most able individuals, does not qualitatively change the implications of the theoretical framework.

To see this, consider the case where the cost of access is $C = \frac{c}{\alpha}$, where c is a constant and α is ability. From inequality 3, it can be seen that the distribution of individuals that enter higher education must satisfy: $\pi\alpha + [w^G - \pi\alpha]G(\underline{t} - m(\alpha)) + \frac{c}{\alpha} \leq w^G$. Defining $q(\alpha) := \pi\alpha + [w^G - \pi\alpha]G(\underline{t} - m(\alpha)) + \frac{c}{\alpha}$, where $q'_\alpha(\alpha) = [1 - G(\underline{t} - m(\alpha))]\pi - m'(\alpha)[w^G - \pi\alpha]g(\underline{t} - m(\alpha)) - \frac{c}{\alpha^2}$. It can be seen that the inclusion of $C = \frac{c}{\alpha}$ to the left-hand side of inequality 3 makes it more likely to be in the case of positive selection into higher education in terms of ability, as it is more likely that the left-hand side of the inequality is a decreasing function of α. Yet, if $m'(\alpha)$ is negative and the public-sector premium is large, negative selection into higher education, in terms of ability, can still occur.

2.3.1 Costly Public-Sector Search and Educated Unemployment

Job search is costly in Ethiopia, particularly public-sector job search [Abebe et al., 2016; Chamberlain, 2015, 2017]. To introduce the opportunity cost of public-sector job search, I introduce a third sector: the skilled private sector. Individuals can remain in the low-skilled sector, as before, at zero cost and receive rW^P, or they can obtain a costly higher education and either: (i) accept high-skilled private-sector employment and receive rW^H; or (ii) search for public-sector employment. In this case, they cannot search for public-sector employment and work simultaneously, hence there is an opportunity cost of searching for public employment. The skilled private-sector consists of homogeneous firms that pay wages equal to marginal productivity, a linear function of ability: $w^H = \pi^H\alpha$, with $\pi^H > \pi^L$. rW^G is the value of searching for public employment. Now, instead of those not selected for public employment earning $\pi\alpha$, they enter unemployment and earn b.

Proposition 2: The most able tertiary-educated individuals select into public-sector job search and have a positive probability of remaining unemployed if there is a positive

correlation between ability and public-service motivation and if the public-sector wage, w^G, is set sufficiently large relative to the outside options (unemployment utility and the private-sector wages).

Proof:

Individuals select into higher education if;

$$\pi\alpha \leq \pi^H\alpha - C \tag{4}$$

Or if:

$$\pi\alpha \leq [1 - G(\underline{t} - \psi)]w^G + G(\underline{t} - \psi)b - C \tag{5}$$

Where inequality 5 is similar to inequality 2, replacing $G(\underline{t} - \psi)\pi\alpha$ with $G(\underline{t} - \psi)b$ on the right-hand side. Inequality 4 implies positive selection into higher education based on ability and, following similar reasoning to above, based on inequality 5, there can be positive selection into higher education only if $m'(\alpha) > 0$ – positive correlation between public-service motivation and ability – and if $[w^G - b] > 0$. Again, if π is low, positive selection into higher eduction based on ability is more likely occur. Hence, if w^G is large relative to both b and π, and there is positive correlation between ability and public-service motivation, the most able individuals will select into higher education. Extending the framework to allow the cost of access to be decreasing in ability does not change the qualitative implications of the theoretical framework and makes it more likely to be in the case of positive selection into higher education in terms of ability.

Among the pool of individuals with a higher education, individuals select into the *public sector* if:

$$\pi^H\alpha \leq [1 - G(\underline{t} - \psi)]w^G + G(\underline{t} - \psi)b \tag{6}$$

Which, inserting $\psi = m(\alpha)$, can be written:

$$\pi^H\alpha + G(\underline{t} - m(\alpha))[w^G - b] \leq w^G \tag{7}$$

Let $h(\alpha) := \pi^H\alpha + G(\underline{t} - m(\alpha))[w^G - b]$, then, similar to above, $h'_\alpha(\alpha) < 0$ and positive selection into public-sector job search, in terms of ability, occurs only if $m'(\alpha) > 0$, if $[w^G - b] > 0$ and if π^H is sufficiently small (the high-skilled private sector is unproductive). Under these conditions, the most able individuals in the high-skilled sector select into public-sector search, with a positive probability of unemployment.

2.4 Identification

The empirical analysis compares individuals in 'exposed' cohorts to the 'unexposed' cohorts, across districts with public universities ('exposed districts') and districts without public universities ('unexposed districts'). Cohorts in 'unexposed' cohorts and 'unexposed' districts are, strictly speaking, 'less exposed', but I use the term 'unexposed' for simplicity.

The basic idea behind the identification strategy is illustrated in figure 5. This section provides further details on the definitions of each of the groups. The LFS estimates rely on $[(1) - (2)] - [(3) - (4)]$. Where $[(1) - (2)]$ captures the effect of a reduction in distance to the nearest university, but also additional differences between the two types of districts ('exposed' districts and 'unexposed' districts) in terms of their labour markets, geographical characteristics, and culture. 'Netting out' $[(3) - (4)]$ from $[(1) - (2)]$ controls for these differences, using the 'unexposed' cohort as the counterfactual group. The ECSS estimates rely on $[(5) - (6)] - [(7) - (8)]$ in figure 5.

The identification strategy uses margins of variation around cohort-geographical pairs, reducing bias associated with differential aggregate conditions, differential human-capital investments at the time of deciding whether to enter higher education and differences in geography, culture, and labour markets. In addition to this straightforward comparison, I further control for cohort fixed effects to fully control for aggregate conditions and potential experience, district fixed effects to control for market conditions at the local level, and individual-level controls to capture differences in individual productivity or tastes, detailed below.

2.4.1 'Exposed' and 'Unexposed' Districts

I define 'exposed' districts as those with a public university within their borders and 'unexposed' districts as those without a public university within their borders. The justification for this is driven by: (i) the limited mobility between districts (see section 2.2.5), (ii) the fact that districts are relatively large units with their own political and administrative structures (representing on average 100,000 citizens), (iii) the remoteness of districts, which are often several hours away from the nearest large urban area (the average is 7 hours) due to the large land mass of Ethiopia and (iv) the additional salience and ease of access of having a public university within the same district and the increased likelihood of interactions with recent graduates.

Nevertheless, to show that the results are not driven by this definition of 'exposed' and 'unexposed' districts, I present the main analyses using three alternative definitions of 'exposed' districts: (i) those less than or equal to 3.1 hours (median) away from the nearest large urban centre; and (ii) those districts less than 40km away (median) from

the nearest public university; and (iii) those that are less than or equal to 3.1 hours away from the nearest large urban centre *and* less than 40km away from the nearest public university.

Figure 2 shows that the universities are geographically spread across different districts. I therefore control for district fixed effects in the main specifications of the regressions to take geographical differences, cultural differences, and local development and labour-market characteristics into account.

2.4.2 'Exposed' and 'Unexposed' Cohorts

2.4.2.1 Labor Force Survey: Education, Employment, and Wages

For the analysis using the LFS, to estimate the effect of the university expansion on education attainment, employment outcomes, and wages, the two main cohorts that I focus on are those aged between 17-20 at the time that the nearest university was established ('exposed') and those aged between 21-24 at the time that the nearest university was established ('unexposed'). The motivation behind this is threefold:

1. The normal age of entry into higher education is 19-20 if an individual follows the standard path of path of progression. The probability of attaining a higher education is lower for individuals that interrupt their studies and do not follow the standard path of progression [Meng and Gregory, 2002].

2. The data shows that there is a noticeable reduction in the probability of having a post-secondary education for those older than 20-years-of-age at the time that the nearest university was established, when individuals have low access costs in terms of distance (reside in the same district as the university). I show this in figures 6 and 17.

3. The normal length of post-secondary courses in universities is 3 years. The application process for universities takes place in the last year of secondary education (normally aged 17-18). Therefore the last age-at-establishment for which the decision to enter higher education is made without being able to observe the labour-market outcomes of the new graduates of the recently established public university is age 17. Those aged 16 may observe the labour-market outcomes of the first set of new graduates from the new local university and respond endogenously. This provides a lower bound for the 'exposed' cohort age band. The upper bound for the 'exposed' cohort is the last age at which the potential entrant can enter higher education at the normal age of entry, aged 20. The first age in the 'unexposed' group therefore is 21. The upper bound for the 'unexposed' band is selected to ensure the two groups are balanced in terms of age range.

To provide evidence that the specific groupings of the cohorts are not driving the core results, I show the main regressions from the LFS comparing only those aged 20 at the time that the nearest university was established ('exposed') and those aged 21 at the time the nearest university was established ('unexposed'). In addition, to alleviate concerns that the results are being driven by the establishment of the university affecting the labour market outcomes of those aged 21-24 at the time of establishment, I run the main analyses using those aged 17-20 as the 'exposed' cohort and using those aged 25-28 as the 'unexposed' cohort.

Since there may still be differences within these definitions of cohorts, in terms of age, so that differences in human capital accumulation could bias the estimates, I control for cohort-of-birth fixed effects in the main specifications of the regressions, which control for all functions of age, and therefore potential experience.

2.4.2.2 Ethiopian Civil Servants Survey

For the analysis with the Ethiopian Civil Servants Survey, to estimate the effect of the university expansion on the characteristics and performance of public officials and therefore measures of public-sector productivity, I define the 'exposed' cohort as civil servants entering the service between 0-4 years *after* the university is able to produce new *graduates*. These are the set of public officials that enter the civil service when there are more graduates on the labour market, due to the establishment of the new university in the district and, as I show in section 2.5, these public officials will also face a larger public-sector wage premium. I compare the characteristics of the 'exposed' cohorts of bureaucrats to those entering the service between 1-4 years *before* the university is able to produce new graduates ('unexposed'). The new university is able to produce new graduates from 3 years after establishment. The 4-year range is used to avoid including individuals in the 'exposed' cohort that could have observed the labour-market returns from the first graduates of the newly established local university before deciding whether to enter higher education or not.

Since there may still be differences between individuals within these cohorts, in terms of experience in the service and potential work experience outside of the service, I control for fixed effects for the cohort-of-entry into the civil service and age and age-squared in the regressions.

2.4.3 Identifying Variation

I use the variation in age-at-exposure and distance from the nearest public university to identify the effect of the university expansion on education choices, in the spirit of Card and Lemieux [2001], Duflo [2001], and Atkin [2016]. Figure 6 shows the prob-

ability of attaining a post-secondary-level education for different cohorts in terms of their age at the time that their nearest university was established. In districts with a public university (left), where access costs in terms of distance are low, there is a noticeable increase in the proportion of individuals with a post-secondary education for the 'exposed' cohort (aged 17-20 at the time that the nearest university was established), relative to the 'unexposed' cohort (aged 21-24 at establishment). This increase is much less apparent in districts without a public university (right), where access costs in terms of distance matter, suggesting the complementarity between distance and age as constraints to higher education. Figure 17 shows this with finer two-year age bands, with similar conclusions.

To demonstrate that the interaction between distance and age-at-establishment determines the probability of attaining a post-secondary education with *continuous* variation in distance and age-at-establishment, figure 7 shows a kernel-weighted local polynomial plot of the unconditional probability of having a post-secondary education for each decile-of-distance and age-at-establishment pair. I plot those aged between 17 and 26 at the time that the nearest university was established for presentation purposes. The figure shows that the probability of having a post-secondary education increases in *both* age-at-establishment *and* distance *interactively*. For those furthest away from the nearest public university, the probability of attaining a post-secondary education is similar for ages at establishment. As distance decreases, there is more take-up of higher education, with a greater response from younger cohorts. This is the interaction that the identification strategy in this paper exploits.

Figure 18 plots the probability of having a post-secondary education for the different cohorts for the most remote districts – those at least 4.4 hours away from the nearest urban centre (the median for the 'unexposed' districts). There is no increase in the probability of attaining a post-secondary-level of education when moving from the 'unexposed' to 'exposed' cohorts. This further highlights the complementarity between distance and age in determining the response to the expansion in public universities.

2.4.4 Validity Checks

Table 1 shows t-tests of mean differences between individuals in districts with a public university and districts without a public university across demographic characteristics and employment outcomes for those aged between 21 and 65 using the Labor Force Survey of 1999. The t-test controls for region fixed effects and the standard errors used for the test are clustered at the district level. The table shows that, in 1999, prior to the rapid expansion in public universities during the 2000s, the populations in the districts that end up with a public university and the districts that do not are

statistically balanced across these dimensions.

Table 2 presents the results from a control experiment, similar to Duflo [2001], which shows the results from the following least-squares regression:

$$\begin{aligned} Y_{i,c,d} =& \alpha + \beta_1 1[University_d] + \beta_2 1[21\_to\_24_{i,c,d}] + \delta 1[University_d] \times 1[21\_to\_24_{i,c,d}] \\ &+ f_c + \phi_d + \Pi' X_{i,c,d} + u_{i,c,d} \end{aligned} \tag{8}$$

Where i is the individual, c is the cohort, d is the district. The sample is restricted to two *'unexposed'* cohorts: those aged 21-24 at the time that the nearest university was established and those aged 25-28. $Y_{i,c,d}$ is the outcome of interest, in this case the probability of holding a post-secondary education (columns 1 and 4), the probability of being employed within the last 12 months (columns 2 and 5) and the log hourly wage (columns 3 and 6). $1[University_d]$ is a binary variable equal to one if the district has a public university. $1[21\_to\_24_{i,c,d}]$ is a binary indicator equal to one if the individual is aged 21-24 at the time that the nearest university was established. f_c are cohort-of-birth fixed effects. $X_{i,c,d}$ includes a set of individual controls: an indicator for whether the respondent is female in all regressions and the broad education levels of individuals in the wage regressions – whether an individual has no formal education, some primary level education, or secondary and above. I combine secondary and higher to avoid the issue of endogenous controls, since having a post-secondary education is also an outcome variable of the policy of higher education expansion. ϕ_d is a set of district fixed effects. Columns (1) to (3) do not condition on fixed effects and individual controls, while columns (4) to (6) include controls.

This control experiment tests whether any systematic trends across districts and cohorts might be driving the results, such as mean reversion. The fact that the coefficient of interest (δ in equation 8) is close in magnitude to zero and not statistically significantly different from zero suggests that this is unlikely to be driving the results.

Figure 8 shows the results of the control experiment for each of the cohorts by their age at the time of establishment, up to aged 40. The right-most bar in the figure, for example, shows the coefficient on the interaction between an indicator for whether the individual was aged 33-36 at the time that the nearest university was established and an indicator if the district of the individual has a public university. The sample includes those aged 33-40 at the time that the nearest university was established, hence the comparison is relative to those aged 37-40 at the time of establishment. The next column shows the effect for those aged 29-32 relative to those aged 33-36 at the time of establishment, and so on. The only coefficient that is significantly different from zero is the comparison between those aged 17-20 at the time that the nearest university was

established and those aged 21-24 (the left-most bar in the figure), exactly the 'natural experiment' of interest in the paper.

To test whether the introduction of a new public university changed the demographic composition of the younger cohorts relative to the older cohorts, and therefore whether the results could be driven by changes in demographics, for example selective migration, rather than education and employment decisions, table 7 shows a series of placebo regressions. The table shows the results from regressions of the form:

$$X_{i,c,d} = \gamma_0 + \gamma_1 1[University_d] + \gamma_2 1[Exposed_{i,c,d}] + \gamma_3 1[University_d] \times 1[Exposed_{i,c,d}] + \rho_r + v_{i,c,d} \quad (9)$$

Where $X_{i,c,d}$ are demographic characteristics, the columns in 7. ρ_r are region fixed effects. Standard errors are clustered at the district level. The sample is restricted to the sample of interest: those aged between 17-24 at the time that the nearest university was established. The coefficient on $1[University_d] \times 1[Exposed_{i,c,d}]$, γ_3, is shown and is the coefficient of interest. If γ_3 is significantly different from zero, it could be possible that endogenous migration flows or demographic changes occurred for the younger cohort when the new university was established. I find no evidence of this.

To check for validity in the Ethiopian Civil Servants Survey, testing for differential selection into the civil service across the cohorts of interest in districts with a public university relative to districts without a public university, I present the results from an OLS regression as per equation 9, where $X_{i,c,d}$ are civil servant characteristics in table 8. 'Exposed' cohorts are those that enter the civil service 0-4 years after the nearest public university is able to graduate students. I find no evidence of differential selection into the civil service, on these demographic characteristics, for younger cohorts relative to older cohorts in districts with a public university relative to districts without a public university.

Table 9 shows the average districts-level characteristics for districts with a public universities and districts without a public university using the census of 2007. Table 10 shows the average demographic characteristics of respondents in districts with a public university and districts without a public university. In both cases, the t-tests of the mean differences across the two types of districts control for region fixed effects. The t-tests displayed in table 9 use robust standard errors and those in table 10 cluster standard errors at the district level. The tables show that, after the mid-2000s, the districts differ along a range of margins. However, the key requirement for identification is that there is no differential variation across the cohorts of interest in districts with a public university, relative to districts without a public university, tested in table 7, described above. Furthermore, I control for district fixed effects in the conditional

regressions and the addition of these controls do not change any of the qualitative results, suggesting that differences among district characteristics are not biasing the results.

2.5 Results

2.5.1 The Effect of the Expansion on Education Attainment

Using the Labour Force Survey of Ethiopia for 2013 and motivated by the identification strategy described in section 4.4, I estimate the effect of the expansion in public universities on education attainment, by running the following least-squares regression:

$$1[Post\,Secondary]_{i,c,d} = \alpha + \beta_1 1[University_d] + \beta_2 1[Exposed_{i,c,d}] + \delta\, \mathbf{1[University_d] \times 1[Exposed_{i,c,d}]} + f_c + \phi_d + \Pi' X_{i,c,d} + u_{i,c,d} \quad (10)$$

i represents the individual worker; c the cohort; d the district. $1[Post\,Secondary]_{i,c,d}$ is an indicator equal to one if the individual has a post-secondary education. $1[University_d]$ is an indicator equal to one if the individual resides in a district with a public university. $1[Exposed_{i,c,d}]$ is an indicator equal to one if the individual is aged 17-20 at the time that the nearest university was established, the 'exposed' cohort in the LFS analysis, motivated in section 4.4. The 'unexposed' cohort is those aged 21-24 at the time of establishment and therefore the sample is restricted only to those aged 17-24 at the time that the nearest university was established. f_c are cohort-of-birth fixed effects, which control for aggregate cohort effects and all functions of age or potential experience. $X_{i,c,d}$ includes an indicator for whether the respondent is female. ϕ_d is a set of district fixed effects. I do not control for whether the individual is a migrant in the regressions due to the possibility of endogenous migration related to education opportunities and wages. I show that there is no evidence of endogenous migration in response to the establishment of the new public university in section 4.4.

β_1 identifies the conditional average effect on $1[Post\,Secondary]_{i,c,d}$ of residing in a district with a public university for the 'unexposed' cohort, capturing average differences between districts with and districts without public universities.[17] β_2 identifies the conditional average effect on $1[Post\,Secondary]_{i,c,d}$ of being in the 'exposed' cohort relative to the 'unexposed' cohort in the districts without a public university, capturing average differences between the two cohorts. δ is the parameter of interest. δ captures the causal effect of the expansion in public universities on $1[Post\,Secondary]_{i,c,d}$ if the

[17] $1[University_d]$ will drop out of the regressions when I control for district fixed effects.

identification strategy is valid (that the differences between the two types of cohorts in the districts without a public university provide a valid counterfactual for the difference between the two cohorts in districts with a public university, explored in section 4.4), conditional on the controls for human capital, tastes, and local labour-market characteristics. Standard errors are clustered at the district level in all regressions.

The results are presented in table 3, which shows the estimates from equation 10 unconditional on fixed effects or additional controls in column 1; with region fixed effects (column 2), cohort-of-birth fixed effects (column 3), distance-to-nearest-university controls (column 4), district fixed effects (column 5), and individual controls (column 6).[18] Column (6) shows the full specification described in equation 10.

Throughout all of these specifications, there is significant evidence that the expansion in public universities led to an increase in tertiary education attainment. The size of the effect is large in economic magnitude, suggesting an increase of 3.4 percentage points once all controls are included (column 6), which is 31% of the mean for the 'unexposed' cohort in the districts without a public university. In addition, the fact that the estimate of δ is stable across all specifications provides further justification for the validity of the identification strategy: the inclusion of controls for district characteristics and controls for individual characteristics do not qualitatively change the results and δ is statistically significant at usual levels across all specifications.

To investigate whether individuals make education decisions based on the public-sector wage premium, as suggested in section 4.3, I show estimates of δ from a least-squares regression as per equation 10 for each percentile of the expected public-sector wage premium at the time that individuals make the decision of whether to enter higher education in figure 9 (top-left). The expected public-sector wage premium is calculated as the difference in average public-sector wages and private-sector wages for those aged 25 and older at the time that the nearest public university was established multiplied by the proportion of public employment for the post-secondary educated for the same age group. The wages of those aged 25 and older are used since they are observable to prospective higher education entrants at the time of entry in university and since they are less likely to be affected by the introduction of the new public university. The figure shows that the expansion in higher education increases education attainment *only* when the public-sector wage premium is high. This finding corresponds to the conceptual framework, both propositions 1 and 2, which state that the public-sector wage premium must be large for individuals to select into higher education.

[18]Distance controls include an indicator for whether the district is neighbouring a district with a public university within it, to control for potential spillover effects; and a measure of the distance from the centroid of the district polygon to the university, this is calculated in degrees and then converted to approximate kilometres by multiplying by 110.567. These are absorbed once districted fixed effects are included.

This is true when comparing the just average wage differences (top-right) and when using only the public-sector wage premium of those aged 25-28 at the time that the nearest public university was established to represent a public-sector wage premium that is closer to the premium available to new labour-market entrants (bottom-left). The public-sector wage premium is correlated with the post-secondary wage premium (correlation coefficient of 0.6), but since many post-secondary-educated individuals accept low-wage jobs in the private sector, the interaction of the effect of the university expansion by the post-secondary wage premium (bottom-right) is much less stark.

This set of results demonstrates the importance of the cross-sector comparison – the prospect of public-sector wages and unattractive private-sector wages – for education decisions, consistent with propositions 1 and 2 of section 4.3, since attractive public-sector jobs are largely restricted to those with higher education. To avoid concerns that the expansion of universities was experienced in areas only where the public-sector premium was large, I perform a t-test comparing 'exposed' districts and 'unexposed' districts in terms of the public-sector wage premium for all individuals and find no significant difference; I conclude the same with a Kolmogorov-Smirnoff test of the distributions of the public-sector wage premium between 'exposed' and 'unexposed; districts and figure 19 plots the distributions.

2.5.2 The Effect of the Expansion on Employment

To estimate the effect of the expansion in public universities on employment, I estimate equation 10 with $1[Employed]$ as the dependent variable, an indicator equal to one if the individual has been employed in any kind of job over the last 12 months.[19] Figure 10 shows the results for the full sample of individuals aged between 17-24 at the time that the nearest university was established in the left bar, for only those with a post-secondary education in the second bar, for only those without a post-secondary education in the third bar, and for those with a post-secondary education and not currently at school in the fourth bar.

The results show that overall employment decreased by 3 percentage points ($p = 0.052$), but this effect is driven by the decrease in the employment rate of high-skilled individuals (6 percentage points, with a p-value of 0.004) and this is not simply because the high skilled are in school. Combined with the fact that the university expansion resulted in more high-skilled individuals on the labour market in areas where the public-sector wage premium was large, this is consistent with relatively attractive and fixed public-sector wages resulting in 'queuing' for public employment, generating educated unemployment, in line with proposition 2 of section 4.3 [Todaro, 1969; Fan and Stark,

[19] $1[Employed]$ is equal to one if the individual responds affirmatively to "Did you engage in any work for pay or profit or family gain for most of the last 12 months?"

2007; Quadrini and Trigari, 2007; Albrecht et al., 2015; Gomes, 2015; Girsberger and Meango, 2017].

2.5.3 The Effect of the Expansion on Wages and Productivity

To estimate the effect of the expansion in public universities on wages, I first run a regression of the log hourly wage on $1[University_d]$, $1[Exposed_{i,c,d}]$, and the interaction between the two indicators. The result of this regression is shown in column (1) of table 4 and provides evidence of a negative effect on wages overall. For the remaining columns of table 4, the regressions take the form:

$$\begin{aligned} &Log\,Hourly\,Wage_{i,c,d} = \alpha + \beta_1 1[University_d] + \beta_2 1[Exposed_{i,c,d}] \\ &+ \delta_1 \mathbf{1[University_d]} \times \mathbf{1[University_d]} + \beta_3 1[Public_{i,c,d}] \\ &+ \beta_4 1[University_d] \times 1[Public_{i,c,d}] + \beta_5 1[Exposed_{i,c,d}] \times 1[Public_{i,c,d}] \\ &+ \delta_2 \mathbf{1[University_d]} \times \mathbf{1[Exposed_{i,c,d}]} \times \mathbf{1[Public_{i,c,d}]} + f_c + \Pi' X_{i,c,d} + \phi_d + u_{i,c,d} \end{aligned} \tag{11}$$

Where $1[Public_{i,c,d}]$ is an indicator equal to one if the individual is a public employee. In the wage regressions, in addition to an indicator for the gender of the respondent, $X_{i,c,d}$ includes indicators for the broad education level of the individual to capture broad differences in human capital – whether an individual has no formal education, some primary level education, or secondary and above. I combine secondary and all above to avoid the issue of endogenous controls, since having a post-secondary education is also an outcome variable of the policy of higher education expansion. There is no effect on secondary education attainment, discussed in section 4.4 and shown in table 7.

In equation 11, δ_1 and δ_2 are the coefficients of interest. δ_1 represents the effect on wages in the *private sector* for the 'exposed' cohort in the 'exposed' district, netting out the average differences across cohorts and the average difference across districts, using the 'unexposed' cohorts and districts as the counterfactuals; and δ_2 represents the marginal effect on wages in the *public sector* for the 'exposed' cohort in the 'exposed' district. The total effect of the university expansion on wages in the public sector is the sum of δ_1 and δ_2.

Column (2) of table 4 shows the results of the regression as per equation 11 without any controls. Column (3) adds the full set of controls; column (4) reports the monthly wages in USD to show the effects on earnings; and column (5) shows the effects on monthly wages in USD for the employed and the non-employed, where the non-employed

are given imputed monthly wages equal to the average private-sector wage in their district for those aged 25-28 at the time that the nearest university was established (the cohort just older than the 'unexposed' cohort). The effects in column (5) are an estimate of the overall effects on earnings, unconditional on employment.

The table shows a significant and negative value for δ_1 across all specifications, evidence that the private-sector wages and earnings responded negatively to the expansion in higher education. Public-sector wages and earnings, however, represented by $\delta_1 + \delta_2$ are around zero in value and not statistically significantly different from zero in columns (2) to (5). The effects on public-sector wages become negative in magnitude when using imputed earnings in column (5) but the total effect on public-sector earnings remains not statistically significantly different from zero. The estimates suggest a 0.13 log-point reduction (12 percentage points) in private-sector wages overall on average, as a result of the expansion in public universities. These estimates imply a 5 USD reduction in monthly earnings for the average private-sector worker, which is 60 USD per year, or 8.5% of the average GDP per capita in Ethiopia in 2016 (707 USD). Using imputed earnings implies a decrease in monthly earnings of 2 USD or 24 USD per year, 3.4% of GDP per capita.

Figure 11 shows the estimates of the effects of the expansion on private-sector wages (red dots and dashed lines) and public-sector wages (blue dots and dashed lines) for each percentile of the expected public-sector premium, similar to figure 9. The private-sector wage results report the coefficient on δ_1 in equation 11. The public-sector wage effects reports the total effect on public-sector wages represented by the sum of δ_1 and δ_2 in equation 11. The figure provides evidence of the constant nature of public sector wages, which do not respond to the shock – the coefficients remain around zero for all districts. The coefficients become slightly more positive in districts where the expected public-sector wage premium is largest, as the population becomes more educated and individuals enter the public sector on a higher civil-service grade. The effect on the private-sector wage becomes negative as the expected public-sector premium increases, in the same districts where the increase in educational attainment is larger as a result of the university expansion.

The results are consistent with positive selection, in terms of ability (productivity), into higher education and therefore into public-sector job search – described in propositions 1 and 2 in section 4.3. As a result, the most productive individuals exit the private sector and do not re-enter – they either remain in non-employment (searching for public-sector jobs) or enter public employment. Because of this, the marginal productivity in the private sector decreases, resulting in lower wages. Further evidence of this selection is provided in the quantile wage regressions shown in table 11, where I run the regression of equation 11 for each quintile of log hourly wages. The negative

effects on private-sector wages are observed in top end of the wage distribution – the wages of the most productive individuals. Propositions 1 and 2 refer to the correlation between public-service motivation and ability, which I present evidence on in sections 2.5.4 and 2.6.2.

Table 12 presents a similar set of regression as equation 11 replacing $1[Public_{i,c,d}]$ with an indicator for whether the individual has a tertiary education (1[Post-secondary]) to show the effects of the expansion on wages by education levels. Column (1) shows the overall effect of the interaction, without breaking up the effects by education level. Column (2) interacts the effects with 1[Post-secondary]. Column (3) adds cohort fixed effects, district fixed effects and individual controls. Column (4) shows the effects on monthly earnings. Note that, since the post-secondary-education premium is of interest $X_{i,c,d}$ includes only an indicator for whether the respondent is female and an indicator for whether the respondent has a post-secondary education rather than the broader set of education-level dummies. The results in table 12 show that individuals without a post-secondary education experience a negative effect on wages, in similar magnitude to the decrease in private-sector wages shown in table 4, while those with a post-secondary education experience no significant change in wages. This increases the returns to higher education. These results are consistent with the effects on the public-sector wage premium: low-skilled individuals make up the majority of the private sector and hence experience the negative effects on wages after the expansion in higher education; high-skilled individuals predominately work in the public sector or enter non-employment after the expansion in higher education and hence experience no change in wages.

2.5.3.1 A Conservative Estimate of the Effects on Private-Sector Productivity

I develop a conservative estimate of private-sector productivity based on wages, allowing for frictions in the labour market such that wages are not exactly equal to productivity. Following Albrecht et al. [2015], I assume that the reservation utility of the individual is captured by their observable characteristics (education level, year of birth, and gender). I assume that that the search parameters for the individual in the local labour market are partially captured by district (local labour market) fixed effects.

To remain conservative, I assume that there remain unobservable search frictions and follow the most conservative stance of Gomes [2017], assuming that 20% of the residual from a regression of the log hourly wage on observable characteristics and district fixed effects is labour productivity. I do this for each 'market' separately, that is: (i) public-sector high-skilled; (ii) public-sector low-skilled; (iii) private-sector high-

skilled; (iv) private-sector low-skilled, where 'high-skilled' refers to those with a post-secondary education or higher and 'low-skilled' refers to those without a post-secondary education. Note that the public sector is primarily high-skilled and the private sector is primarily low-skilled. I do this only for those aged between 17 and 24 at the time that their nearest university was established to be able to apply the identification strategy described in section 4.4 to the residuals and identify the average effects of the university expansion on labour productivity.

I regress the measure of private-sector productivity on an indicator for whether the individual was in the 'exposed' cohort, an indicator for whether the individual is in the 'exposed' district, and an interaction between the two, without the fixed effects and observable characteristics (since the residuals are from a regression in which these are already controlled for). The results are shown in figure 12. The first bar reports the results for high-skilled public-sector workers; the second bar for low-skilled public-sector workers; the third bar for high-skilled private-sector workers; the fourth bar for low-skilled private-sector workers; and the fifth bar for the private sector overall.

The estimates suggest that the private sector experienced a 2 percentage point decrease in productivity based on this conservative estimate, driven by the reduction in productivity of the low-skilled, consistent with propositions 1 and 2 of section 4.3. Using public-sector wages to 'back out' productivity is problematic in this setting as public wages are rule-based and not clearly linked to productivity. I explore the impact of the higher education expansion on public-sector productivity in section 2.5.4.

2.5.4 The Effect of the Expansion on Public-Sector Productivity

Measuring public-sector productivity from wages is ill-advised, considering the rigid and rule-based nature of wages, not necessarily linked to worker productivity [Finan et al., 2017]. For example, I find that the score of the civil servant in her annual performance appraisal explains only 3.5% of the variation in the measure of wages.[20]

I therefore turn to the Ethiopian Civil Servants Survey data, which I helped design and implement for the purpose of this study, to estimate the effects of the expansion in higher education on the characteristics and performance of civil servants entering the service. The ECSS data contains demographic information, measures of public-service motivation [Perry, 1996], measures of effort (hours worked), and civil service performance appraisals for a representative sample of public officials in Ethiopia across five sectors. With the data from the ECSS, I run the following least-squares regression:

[20]The civil service grade is used instead of wages, since wages were not captured in the survey and the civil service grade is mapped on to a matrix of wages.

$$Y_{i,c,d} = \alpha + \beta_1 1[University_d] + \beta_2 1[Exposed\,Graduates_{i,c,d}] + \delta\, \mathbf{1[University_d]} \times \mathbf{1[Exposed\,Graduates_{i,c,d}]} + f_c + \phi_d + \Pi' X_{i,c,d} + u_{i,c,d} \quad (12)$$

$1[University_d]$ is an indicator equal to one if the individual resides in a district with a public university. $1[Exposed\,Graduates_{i,c,d}]$ is an indicator equal to one if the civil servant entered the service 0-4 years after the nearest university started to graduate students, the 'exposed' cohort in the ECSS analysis, motivated in section 4.4. The 'unexposed' cohort is those that entered the service 1-4 years before the university started to graduate students and therefore the sample is restricted only to those that entered the service between 4 years before and 4 years after the nearest university started to produce graduates. f_c are cohort-of-entry-into-the-service fixed effects, which control for experience in the public sector. $X_{i,c,d}$ includes an indicator for whether the respondent is female, an indicator for whether the individual has a post-graduate degree, age and age-squared to account for potential experience outside the civil service. ϕ_d is a set of district fixed effects. δ is the coefficient of interest and represents $[(5)-(6)]-[(7)-(8)]$ in figure 5, without conditioning on cohort fixed effects, districts fixed effects and individual controls.

To estimate the impact of the university expansion on the types of bureaucrats entering into the civil service and their effort and performance, I use the following as my dependent variables $Y_{i,c,d}$, in each column of table 5:

1. The public-service motivation of the bureaucrat, as per the scale in Perry (1996), in z-scores (columns 1 and 5);

2. A measure of the extent to which the bureaucrat sees her mission as aligned to that of the organisation: The stated response to "On what percentage of tasks/projects does your organisation fulfil its mission?" (columns 2 and 6);

3. The stated number of hours worked in a typical week (columns 3 and 7);

4. The performance evaluation score of the civil servant, out of 100. The performance evaluation score is based on a detailed matrix of objective indicators for tasks that the civil servant is expected to complete. "The performance score is based on a weighted average of a matrix of individual tasks and is considered verifiable and subject to appeal from employees – the data also confirms that the performance score is a significant function of visible effort [(hours worked per week)]" [Somani, 2017]. The data is only available for a subset of civil servants (columns 4 and 8).

Columns (1) to (4) of table 5 show the results for each of these outcome variables without individual controls and fixed effects and columns (5) to (8) show the results when conditioning on all controls. I find that the expansion in higher education leads to significantly more motivated civil servants (0.7 standard deviations, column 5, an 11% increase relative to the counterfactual mean of the 'unexposed' cohorts in the 'unexposed' districts) and that these civil servants are significantly more likely to feel that the organisation's mission is aligned to their own (38.4 percentage points, column 6, a 57% increase relative to the counterfactual mean). These civil servants work significantly more hours per week (5.99, column 7, a 14% increase relative to the counterfactual mean), despite earning the same salary, suggesting greater effort. Column 8 shows that these civil servants achieve higher scores in their performance appraisals (3.76 percentage points, column 9, a 4.5% increase relative to the counterfactual mean).

This set of results is in line with propositions 1 and 2, with evidence of more productive and motivated workers in the public sector after the expansion in higher education. The positive effects on both the performance evaluations and public-service motivation provide evidence that these characteristics are positively correlated. I explore this relationship further, for the population as a whole, in section 2.6.2.

2.5.4.1 Management Practices Matter

To investigate which types of public organisations benefitted most from the expansion in higher education, I explore whether management practices related to screening for talent matter. I turn to data on organisational management practices based on the World Management Survey [Bloom and Van Reenen, 2007*a*]. In the spirit of Rasul and Rogger [2016] and Rasul et al. [2018*b*], I group management practices of public organisations into three aggregate practices: 'screening', 'incentives' and 'other'.

The ECSS adapted the method of Bloom and Van Reenen [2007*a*] to cover seven dimensions of management practice: incentives, monitoring, targeting, staffing, staff involvement, roles, and flexibility. Table 13 and table 14 detail each management question, by topic, as well as the 1-5 scoring grid used by the enumerators for each question.

Following Rasul and Rogger [2016] and Rasul et al. [2018*b*], I group incentive-based management practices together, covering the monitoring, targeting and incentive topics. I then group screening-related management practices together, covering staffing and flexibility, since these topics focus on the attraction and retention of talent ("Do you think the management of your Directorate think about attracting talented people to your Directorate and then doing their best to keep them?") and responding to changes in the local environment or local shocks ("How flexible would you say your Directorate is

in terms of responding to new and improved work practices or reforms?"). Organisations that place a heavy weight on staffing and flexibility should be well-placed to identify talent among the new, larger pool of graduates and adapt the organisation's structure and work practices to best suit the available talent on the labour market. Finally, I group other management practices together covering the roles and staff involvement topics, to also control for these dimensions of organisational practices.

Each question was asked to the managers of organisations in a face-to-face interview. The interviews were "double-blind", so managers were not told in advance they were being scored or shown a score grid and enumerators were given no information on the organisation's performance. Z-scores were created for each question, relative to the mean and standard deviation across all managers. Topic z-scores are then the means across different individual questions within topics. I then construct the aggregate 'screening' index as the mean of z-scores across staffing and flexibility. I construct the 'incentives' index as the mean of z-scores across incentives, monitoring and targeting. Similarly, the 'other' index is the mean of z-scores across roles and staff involvement. Organisation averages are used to capture the management practices across all divisions within the organisation.

Figure 20 shows the coefficient on $1[University_d] \times 1[Exposed\,Graduates_{i,c,d}]$, from a regression as per equation 12 with the z-score of public-service motivation as the dependent variable, corresponding to column (7) of table 5, conditional on the full set of controls, but separately for each percentile of management practices related to screening. The figure shows that only those organisations that place a large weight on management practices related to screening are able to attract and employ significantly motivated bureaucrats after the expansion in higher education. Those with the lowest levels of management practices related to screening experience no significant increase in the public-service motivation of staff, while those at the 75th percentile experience a 1 standard-deviation increase in the public-service motivation of staff.

These results suggest that having management practices in place that allow operations to respond to changing local conditions and that emphasise the attraction of talent are essential for public-sector organisations to benefit from the increase in the supply of skilled personnel.

2.5.5 Robustness

To demonstrate that the results are not sensitive to specific definitions of the 'exposed' cohorts or 'exposed' districts, I present several robustness checks. Table 15 shows that all of the main results hold qualitatively when using the most narrow variation between cohorts: comparing those aged 20 at the time that the nearest public university

was established ('exposed') and those aged 21 at the time that the nearest public university was established ('unexposed'). Table 16 shows that all of the main results hold qualitatively when defining the 'exposed' cohort as those aged 17-20 at the time that the nearest university was established but the 'unexposed' cohort as those aged 25-28 at the time that the nearest university was established. This check alleviates concerns that those aged 21-24 at the time that the nearest university was established could be an inappropriate counterfactual (for example, if they are also heavily affected by the introduction of a new public university).

Table 17 shows the main results using the same definitions of cohort exposure as the main analysis but defining 'exposed' districts as those less than the median number of hours away (3.1) from the nearest large urban area, table 18 shows the main results when defining 'exposed' districts as those less than the median number of kilometres (40) away from the nearest public university, and table 19 shows the main results when defining 'exposed' districts as those less than the median number of hours away from the nearest large urban area and those less than the median number of kilometres away from the nearest public university. All of these results suggest that the qualitative conclusions hold, although the statistical significance of the results under alternative definitions of the appropriate 'market' becomes weaker. This is likely due to the fact that the individual's home district is the appropriate definition of the market in this context, with limited mobility and access to public services attached to strict district residency rules, as detailed in section 5.2.

2.6 Alternative Mechanisms

2.6.1 Selection into Higher Education

To check that the increase in the public-service motivation of recruited public officials is indeed due to selection into public-sector job search and public employment, rather than as a result of younger cohorts demonstrating higher levels of public-service motivation, I turn to the World Values Survey data, which contains measures of public-service motivation for individuals in *both* the public sector and private sector. I use data from Wave 5 of the World Values Survey (2007) and follow Cowley and Smith [2014] to analyse the differential intrinsic motivation in the public and private sector associated with the expansion in higher education. Since the World Values Survey is only disaggregated to the regional level, I use the number of public universities available in the region when individuals are within the 'exposed' cohort range (17-20) relative to the number of public universities available when individuals are within the 'unexposed' cohort range (21-24). I run the following regression to estimate the marginal effect of having an additional university available during the 17-20 age range within the region:

$$M_{i,r} = \alpha + \beta_1 Universities\,[17-20]_{i,r} + \beta_2 Universities\,[21-24]_{i,r} + \Psi_r + \mu_c + \Pi' X_{i,r} + u_{i,r} \quad (13)$$

i represents the individual and r the region. $M_{i,r}$ is a measure of intrinsic motivation, $Universities\,[17-20]_{i,r}$ is the number of public universities available to individual when i is aged 17-20 years old; $Universities\,[21-24]_{i,r}$ the number of universities available when i is aged 21-24 years old; Ψ_r region fixed effects; μ_c cohort-of-birth fixed effects; and $X_{i,r}$ includes an indicator for whether the respondent is female. To avoid including endogenous controls, I do not control for the education level of respondents, since the education level is a potential outcome variable of the increase in the expansion of universities. Instead, I control for broad categories of education, including an indicator for whether the respondent has no formal education, primary level, or secondary and above.

Following Cowley and Smith [2014] directly, I use two self-reported measures of intrinsic motivation:

1. Those that respond 'doing an important job' to "Regardless of whether you're actually looking for a job, which one would you, personally, place first if you were looking for a job?"

2. Those that respond 'Very much like me' or 'Like me' to "Using this card, would you please indicate for each description whether that person is very much like you, like you, somewhat like you, not like you, or not at all like you. It is important to this person to help the people nearby; to care for their well-being."

The first measure therefore relates more directly to intrinsic motivation that is general to all types of activities, while the second measure is related more to public-service motivation [Cowley and Smith, 2014]. The results are reported in table 20. Columns (1) to (3) display the results when the first measure of intrinsic motivation is the dependent variable. Columns (4) to (6) display the results for the second measure. The first measure has very little variation in the entire sample, with only 5% of respondents responding affirmatively. To examine the differential impact on private-sector and public-sector workers, I interact $Universities\,[17-20]_{i,r}$ and $Universities\,[21-24]_{i,r}$ with an indicator for whether the individual works in the public sector for each measure of intrinsic motivation (columns 3 and 6).

The results suggest that there is no systematic change in public-service motivation across cohorts as the number of public universities expand (columns 1 to 2 and columns

4 to 5). The results are consistent with differential selection into the public sector by public-service motivation and an increase in the public-service motivation of *public-sector workers* as a result of the expansion in higher education. This is exactly in line with proposition 2 in section 4.3 and the findings discussed in section 2.5.4.

2.6.2 Public-Service Motivation and Ability

Since the implications of the conceptual framework in section 4.3 rely on the correlation between public-service motivation and ability, I explore this relationship in this section. Public-service motivation reduces the cost (disutility) of public-sector job search, hence those with the highest public-service motivation will have a higher probability of receiving the attractive public wage. If public-service motivation is positively correlated with ability, then this selection into public-sector job search will also mean that the most able (productive) individuals have a higher probability of public employment.

Firstly, I turn to the existing literature, which finds no evidence of negative correlation [Ashraf et al., 2016; Dal Bó et al., 2013] and evidence of positive correlation [Ashraf, Bandiera and Jack, 2014]. Secondly, I turn to the World Values Survey data once again to measure the correlation between public-service motivation (the second measure in section 2.6.1 above) and proxies for ability in the population. The two proxies that I use are: (i) the highest education level of the respondent (no formal education, some primary level education, some secondary level education, complete secondary level education, and upper secondary or higher) and (ii) a measure of the cognitive nature of the tasks that the individual performs at work.[21] Since the measure of public-service motivation is binary, I report the results from a regression with the public-service motivation measure as the dependent variable rather than the correlation coefficient.

I run a least-squares regression of the measure of public-service motivation on the set of binary indicators for the highest education level of the individual, conditional on year-of-birth fixed effects, gender, and region fixed effects, with robust standard errors. The results show that the *only* the coefficients on whether the individual has a complete secondary level education (a coefficient of 0.11 and p-value of 0.064) and whether the individual has an upper-secondary level education or higher (a coefficient of 0.18 and p-value of 0.002) demonstrate significantly higher levels of public-service motivation than those without formal education. A least-squares regression of the measure of public-service motivation on the variable measuring the cognitive nature of the tasks that the respondent performs at work, conditional on the same set of controls, gives a

[21] The second proxy is the response to: 'Are the tasks you perform at work mostly manual or mostly cognitive? If you do not work currently, characterize your major work in the past. Use this scale where 1 means mostly manual tasks and 10 means mostly cognitive tasks.'

positive and significant coefficient (a coefficient of 0.03 and a p-value of 0.00). Both of these results provide further suggestive evidence of a positive correlation between public-service motivation and cognitive ability in the population.

2.6.3 Financial Constraints

An alternative explanation for the increase in education attainment as a result of the expansion in higher education is related to financial constraints: it is plausible that individuals were previously unable to access higher education due to financial constraints rather than rational calculations based on their ability and expected returns of undertaking a post-secondary education. This would lead to selection into higher education (and therefore the public sector) based on wealth rather than on ability or public service motivation, which may be positively or negatively correlated with wealth.

To analyse whether the cohorts that entered higher education after the expansion were more financially constrained than previous cohorts, I use the household Living Standards Measurement Survey (LSMS) and run the same regression as equation 10 with dependent variables as proxies for individual wealth. To proxy for household wealth, I use the following measures: an indicator for whether the mother has completed at least primary education; an indicator for whether the mother has completed at least secondary-level education; an indicator for whether the father has at least a primary-level education; an indicator for whether the father has at least a secondary-level education; and the total number of listed assets that the household reports in the LSMS. As before, I control for district fixed effects, cohort-of-birth fixed effects, and an indicator for whether the individual is female. The definitions of 'exposed' cohorts are as before – those aged between 17 and 20 at the time that their nearest university was established – and 'unexposed' cohorts are those aged between 21 and 24 at the time of establishment. 'Exposed' districts are those districts with a public university and 'unexposed' districts are those without. The sample is restricted to those aged between 17 and 24 at the time their nearest university was established and those aged between 21 and 65 at the time of the LSMS survey.

Table 21 presents the results for the coefficient on $1[University_d] \times 1[Exposed_{i,c,d}]$ for the full sample (columns 1 to 5) and for the sample with a post-secondary education (columns 6 to 10). The estimates suggest no significant difference in household wealth related to the expansion in public universities in any of the cases.

2.6.4 Effects on Consumption

Finally, since the non-employment rate increases and since alternative employment is potentially available (although I find no increase into informal employment), I investi-

gate whether there are effects on reported consumption as a result of the higher education expansion using the LSMS data. I run the same regression as equation 10 with the reported monthly household expenditure from the LSMS data as the dependent variable. The results are reported in table 22 which shows the estimates of the coefficient on $1[University_d] \times 1[Exposed_{i,c,d}]$ for the full sample of individuals within the 'exposed' cohort and 'unexposed' cohort range in column (1), for those employed over the last 12 months in column (2), for those not employed over the last 12 months in column (3), for all post-secondary-educated individuals in column (4), for post-secondary-educated individuals employed over the last 12 months in column (5), and for post-secondary-educated individuals not employed over the last 12 months in column (6).

The findings suggest that the post-secondary-educated individuals experienced a significant reduction in consumption as a result of the expansion in higher education, consistent with the results on employment and wages in section 2.5. Furthermore, the employed and non-employed post-secondary educated individuals both experienced a significant reduction in consumption, suggesting that individuals did not have access to alternative sources of income.

2.7 Conclusion

This paper shows that public-sector hiring and wage policies can determine the effectiveness of policies that aim to increase access to higher education. Public-sector contracting policies and higher-eduction policy should, therefore, be designed hand-in-hand. This conclusion is of particular importance in developing countries where a large majority of tertiary-educated individuals work in the public sector and where attractive private-sector jobs for high-skilled labour are rare.

Attractive public-sector jobs that are reserved for the most-educated in society can determine the entire distribution of skills in the economy. Since the number of public-sector jobs and the level of the public-sector wage (as well as other benefits) jointly impact the returns to higher education, public-sector contracting policies determine whether individuals will take up a higher level of education or not. Furthermore, it is likely that certain types of degree subjects are more valued by public-sector organisations than others (for example, public administration), hence public-sector contracting policies can also determine what subjects individuals choose to study. This means that public-sector contracting policies have the potential to determine the distribution of skills in the economy along both dimensions, with implications for future growth potential [Acemoglu and Zilibotti, 2001; Banerjee and Duflo, 2005; Hsieh and Klenow, 2009].

This paper provides evidence that, given attractive and rigid public-sector wages

for tertiary-educated individuals, expansions in higher education can lead to negative effects on private-sector productivity as the most productive individuals leave the private sector to enter higher education and public-sector job search. I find evidence that public-sector productivity increases. While these results provide evidence of a trade-off, it is important to point out the relatively short-term nature of these results. The expansion in public universities in Ethiopia occurred during the 2000s and the data used to estimate the effects are from 2013 and 2016. The potential long-run implications on private-sector productivity are necessary to quantify to fully estimate the welfare implications of the higher education policy.

Theory Appendix

Proposition 1 states that there is negative selection, in terms of ability (α), into higher education if ability (α) and public-service motivation (ψ) are independent. If α and ψ are independently distributed in the population and if w^G is sufficiently large, there exists a function with $\overline{\alpha}(\psi) = a + \kappa\psi$, where $a > 0$ and $\kappa > 0$ and inverse $\overline{\psi} = \frac{-a}{\kappa} + \frac{1}{\kappa}\alpha$ that describe the locus of all types indifferent between entering higher education and private-sector employment directly. Given ψ, all those with $\tilde{\alpha} < \overline{\alpha}$ strictly want to enter higher education and all those with $\tilde{\alpha} > \overline{\alpha}$ strictly prefer to enter private-sector employment directly. Given α, all those with $\tilde{\psi} > \overline{\psi}$ strictly prefer to enter higher education and all those with $\tilde{\psi} < \overline{\psi}$ strictly prefer private-sector employment directly.

Following inequality 2, the infinite pairs (α, ψ) that are indifferent between entering higher education and not entering higher education satisfy:

$$\pi\alpha = [1 - G(\underline{t} - \psi)]w^G + G(\underline{t} - \psi)\pi\alpha - C \tag{14}$$

The implicit function theorem ensures that a continuous function $\overline{\alpha}(\psi)$ exists, which maps the public-service motivation parameter to the highest-ability type that is indifferent between entering higher education and not.

Rewriting equation 14 as the net gains from entering higher education:

$$[1 - G(\underline{t} - \psi)][w^G - \pi\alpha] - C = 0 \tag{15}$$

From this, $\frac{d\alpha}{d\psi} = \frac{g(\underline{t}-\psi)[w^G - \pi\alpha]}{\pi(1-G(\underline{t}-\psi))}$ can be derived, where $\frac{d\alpha}{d\psi} > 0$ if there is a positive wage public-sector premium $[w^G - \pi\alpha] > 0$. To find the intercept of $\overline{\alpha}(\psi)$, consider the net returns when $\psi = 0$. This is positive if $w^G > \frac{C}{1-G(\underline{t})}$, i.e. if the public-sector wage (w^G) is sufficiently large such that all types face a positive expected return to public-sector job search.

In this case $\overline{\alpha}(\psi)$ has a positive slope and positive intercept. Invertibility yields

$\overline{\psi} = \frac{-a}{\kappa} + \frac{1}{\kappa}\alpha$, where $\kappa = \frac{g(\underline{t}-\psi)[w^G - \pi\alpha]}{\pi(1-G(\underline{t}-\psi))}$.

To see the direction of selection into higher education in terms of ability and public-service motivation, consider the indifferent type (α', ψ'). The claim is that the least able enter higher education, i.e., that $(\tilde{\alpha}, \psi')$ strictly prefers to enter if $\tilde{\alpha} < \alpha'$. Suppose this is not true, then the net gains from entering higher education must be negative for this type:

$$[1 - G(\underline{t} - \psi')][w^G - \pi\tilde{\alpha}] - C < 0$$

Since $\tilde{\alpha} < \alpha'$ and since the net gains are decreasing in α, it must also be true that:

$$[1 - G(\underline{t} - \psi')][w^G - \pi\alpha'] - C < 0$$

But this is a contradiction since type (α', ψ') is indifferent. Similar logic shows that all those with $\tilde{\alpha} > \overline{\alpha}$ strictly prefer to enter private-sector employment directly and that all those with $\tilde{\psi} > \overline{\psi}$ strictly want to enter higher education, while all those with $\tilde{\psi} < \overline{\psi}$ strictly prefer private-sector employment directly.

Table 1: Characteristics of Individuals in Districts with and without a Public University, Labor Force Survey 1999

	(1)		(2)		T-test
	No Public University		Public University		Difference
Variable	N/[Clusters]	Mean/SE	N/[Clusters]	Mean/SE	(1)-(2)
Age	108427	37.28	11290	36.63	0.65**
	[344]	(0.11)	[30]	(0.35)	
1[Female]	108427	0.53	11290	0.52	0.01
	[344]	(0.00)	[30]	(0.01)	
1[Married]	108427	0.77	11290	0.75	0.02
	[344]	(0.00)	[30]	(0.02)	
1[Migrate for work]	108427	0.01	11290	0.01	-0.00
	[344]	(0.00)	[30]	(0.00)	
1[Migrate for education]	108427	0.00	11290	0.00	-0.00
	[344]	(0.00)	[30]	(0.00)	
1[High-school education]	108380	0.03	11274	0.05	-0.02
	[344]	(0.00)	[30]	(0.01)	
1[Post-secondary education]	108380	0.01	11274	0.01	-0.00
	[344]	(0.00)	[30]	(0.00)	
1[Currently employed]	108413	0.81	11284	0.79	0.02
	[344]	(0.01)	[30]	(0.02)	
Hours of work, all jobs last week	73182	24.30	7269	27.02	-2.73
	[344]	(0.34)	[30]	(1.36)	
1[Public sector employee]	85971	0.04	8265	0.07	-0.03
	[344]	(0.00)	[30]	(0.02)	
F-test of joint significance (F-stat)					1.13

Notes: * p<0.1, ** p<0.05, *** p<0.01. Standard errors in brackets. Clustered standard errors are reported. Standard errors are clustered at the district level. The unit of observation is the individual. The sample is restricted to those aged between 21 and 65 only. 1[Migrate for work] is equal to one if the individuals has migrated for employment or job search. 1[Migrate for education] is equal to one if the individual migrated for education. 1[High-school education] is equal to one if the highest level of education attained is secondary. 1[Post-secondary education] is equal to one if the individual attained a tertiary education. 1[Currently employed] is equal to one if the respondent is currently employed, relative to unemployed or inactive. Hours of work, all jobs last week is the reported number of hours that the respondent worked across all jobs in the last week. 1[Public sector employee] is equal to one if the individual is working in the public sector, conditional on the individual being employed. The 1999 Labor Force Survey of Ethiopia does not contain information on wages, unlike the 2013 wave. T-tests are conducted controlling for region fixed effects. Figures are rounded to two significant figures.

Table 2: Control Experiment: Using Unexposed Cohorts Only

	Unconditional			Conditional		
	(1)	(2)	(3)	(4)	(5)	(6)
	Post-Secondary Education	Employed Last 12 Months	Log Hourly Wage	Post-Secondary Education	Employed Last 12 Months	Log Hourly Wage
Exposed District	0.11***	-0.080***	1.52***	0.12***	-0.24***	1.21***
	(0.023)	(0.016)	(0.16)	(0.0063)	(0.0062)	(0.040)
Exposed Cohort	0.035***	-0.016***	0.034	0.0031	-0.012	0.017
	(0.0046)	(0.0055)	(0.061)	(0.016)	(0.014)	(0.079)
Exposed District x Exposed Cohort	0.0067	-0.013	-0.026	0.0034	-0.016	0.047
	(0.012)	(0.012)	(0.098)	(0.011)	(0.011)	(0.039)
Cohort of Birth FE	No	No	No	Yes	Yes	Yes
Individual controls	No	No	No	Yes	Yes	Yes
District FE	No	No	No	Yes	Yes	Yes
Adjusted R-Squared	0.027	0.012	0.15	0.12	0.15	0.69
Observations	23894	23894	9190	23894	23894	9190
Clusters	596	596	580	596	596	580

Notes: * p<0.1, ** p<0.05, *** p<0.01. OLS Estimates. Clustered standard errors in brackets. Standard errors are clustered at the district level. The unit of observation is the respondent. The sample is restricted to those aged between 21 and 28 at the time that the nearest university was established and those aged between 21 and 65 years of age at the time of the survey. Exposed Cohort is an indicator equal to one if the individual is aged 21-24 at the time that the nearest university is established. Exposed District is equal to one if the respondent lives in a district with a public university. The dependent variable in columns (1) and (4) is an indicator for whether the respondent holds a post-secondary education. The dependent variable in columns (2) and (5) is an indicator for whether the individual has been employed in the last 12 months. The dependent variable in columns (3) and (6) is the log hourly wage. Columns (4) to (6) include cohort fixed effects, and district fixed effects. Individual controls in column (4) and (5) include an indicator for whether the respondent is female. Individual controls in column (6) include an indicator for whether the respondent is female and an indicator for whether the individual has no formal education, primary education, or at least secondary education. Figures are rounded to two significant figures.

Table 3: Higher Education Attainment (Post-Secondary Education Probability)

	(1) Baseline	(2) Region FE	(3) Cohort FE	(4) Distance Controls	(5) District FE	(6) Individual Controls
Exposed Cohort	0.012	0.015**	-0.018*	-0.018	-0.022	-0.023
	(0.0084)	(0.0072)	(0.011)	(0.011)	(0.015)	(0.015)
Exposed District x Exposed Cohort	0.035**	0.034***	0.027**	0.027**	0.033***	0.034***
	(0.014)	(0.013)	(0.012)	(0.012)	(0.011)	(0.011)
Cohort FE	No	No	Yes	Yes	Yes	Yes
Region FE	No	Yes	Yes	Yes	No	No
Distance controls	No	No	No	Yes	No	No
District FE	No	No	No	No	Yes	Yes
Individual controls	No	No	No	No	No	Yes
Adjusted R-Squared	0.030	0.040	0.047	0.048	0.099	0.11
Observations	24853	24853	24853	24853	24853	24853
Clusters	596	596	596	596	596	596

Notes: * p<0.1, ** p<0.05, *** p<0.01. Standard errors in parentheses. Standard errors clustered at the district level. The unit of observation is the individual. Exposed Cohort is an indicator equal to one if the individual is aged 17-20 at the time that the nearest university was established, Exposed District is an indicator equal to one if the respondent is in a district with a public university. The sample is restricted to those aged 17-24 at the time that the nearest university was established. The dependent variable is a binary indicator for whether the respondent has a post-secondary-level education. Region fixed effects control for the overarching regional administration under which the district operates. Cohort fixed effects control for all functions of age. Distance controls include an indicator for whether the district is neighbouring a district with a public university, to control for potential spillover effects; and a measure of the distance from the centroid of the district polygon to the university, this is calculated in degrees and then converted to approximate kilometres by multiplying by 110.567. Individual controls include an indicator for whether the respondent is female. Figures are rounded to two significant figures.

Table 4: Log Hourly Wage, Public Sector and Private Sector

	(1) All Sectors	(2) Public & Private	(3) All Controls	(4) Monthly Earnings USD	(5) Imputed Wage Effects
Exposed Cohort	-0.027	0.020	0.015	1.92	0.90
	(0.037)	(0.020)	(0.038)	(1.69)	(1.05)
Exposed District x Exposed Cohort	-0.14**	-0.18**	-0.13***	-5.27***	-2.16**
	(0.061)	(0.072)	(0.042)	(1.25)	(1.05)
1[Public Sector]		2.02***	1.27***	27.2***	25.6***
		(0.057)	(0.092)	(2.51)	(2.47)
Exposed District x 1[Public Sector]		-0.88***	-0.55***	-2.48	-0.25
		(0.11)	(0.11)	(3.28)	(4.16)
Exposed Cohort x 1[Public Sector]		-0.0097	0.012	-2.12	-3.27**
		(0.032)	(0.037)	(1.48)	(1.52)
Exposed District x Exposed Cohort x 1[Public Sector]		0.099	0.12*	3.99	-2.75
		(0.083)	(0.063)	(3.23)	(5.57)
Cohort of Birth FE	No	No	Yes	Yes	Yes
Individual controls	No	No	Yes	Yes	Yes
District FE	No	No	Yes	Yes	Yes
Mean Hourly Wage (USD)	0.18	0.18	0.18	0.18	0.18
Mean Monthly Wage (USD)	115.6	115.6	115.6	115.6	115.6
Adjusted R-Squared	0.12	0.58	0.72	0.55	0.34
Observations	10136	10136	10136	10136	24853
Clusters	581	581	581	581	596

Notes: * p<0.1, ** p<0.05, *** p<0.01. Standard errors in parentheses. Standard errors clustered at the district level. The unit of observation is the respondent. The sample is restricted to those aged between 17 and 24 at the time that the nearest university was established and those aged between 21 and 65 years of age at the time of the survey. The dependent variable is the log hourly in columns (1) to (3), the nominal monthly earnings in USD in column (4), and the imputed monthly earnings in USD in column (5), which is the nominal monthly earnings for those employed and the average private wage for those aged 25-28 in the district at the time the nearest university was established. Individual controls include an indicator for whether the respondent is female and an indicator for whether the individual has no formal education, primary education, or at least secondary education. Exposed Cohort is equal to one if the individual is aged 17-20 at the time that the nearest university was established. Exposed District is equal to one if the individual lives in a district with a public university. 1[Public Sector] is equal to one if the individual works in the public sector. Figures are rounded to two significant figures.

Table 5: Bureaucrat Characteristics and Performance

	Unconditional					Conditional		
	(1) Public Service Motivation	(2) Mission Alignment	(3) Hours worked typical week	(4) Performance evaluation score	(5) Public Service Motivation	(6) Mission Alignment	(7) Hours worked typical week	(8) Performance evaluation score
Exposed Cohort	0.020	-4.32	-2.37*	-3.20**	0.016	-17.0**	1.38	12.5***
	(0.081)	(2.95)	(1.26)	(1.42)	(0.24)	(7.44)	(2.79)	(2.89)
Exposed Cohort x Exposed District	0.58***	26.2***	8.62***	5.95***	0.70***	38.4***	5.99**	3.76*
	(0.13)	(9.54)	(2.80)	(1.42)	(0.12)	(5.66)	(2.38)	(2.01)
Cohort FE	No	No	No	No	Yes	Yes	Yes	Yes
Individual controls	No	No	No	No	Yes	Yes	Yes	Yes
District FE	No	No	No	No	Yes	Yes	Yes	Yes
Adjusted R-Squared	-0.0022	0.0041	0.0069	0.017	0.14	0.053	0.18	0.46
Observations	331	358	358	144	331	358	358	144
Clusters	64	65	65	32	64	65	65	32

Notes: * p<0.1, ** p<0.05, *** p<0.01. Standard errors in parentheses. Standard errors clustered at the district level. The unit of observation is the respondent. Exposed Cohort is an indicator equal to one if the civil servant entered the service 0-4 years after the nearest university started to graduate students. Exposed District is equal to one if the civil servant works in a district with a public university. The sample is restricted to employees entering the civil service between 4 years before and 4 years after the time that their nearest university started to graduate students and those between 21 and 65 years of age. The dependent variable in column (1) and (5) is is the z-score of the Public Service Motivation Scale as per Perry (1996); in column (2) and (6) is the response to "On what percentage of tasks/projects does your organisation fulfil its mission?"; in column (3) and (7) is the stated response to "What is the actual number of hours you work in the civil service in a typical week?"; in columns (4) and (8) is the z-score of the performance-evaluation score of the civil servant. Columns (5) to (8) include a range of controls: cohort-of-entry-into-the-service fixed effects, an indicator for whether the individual has a post-graduate degree, an indicator for whether the individual is female, and district fixed-effects. Figures are rounded to two significant figures.

Table 6: Details of Public Universities

University	Region	Academic Year Established	Year of First Graduates	Special Purpose University	Accredited Private University in District
Adama University	Oromiya	2006	2009	No	Yes
Addis Ababa University	Addis Ababa	1950	1953	No	Yes
Adigrat University	Tigray	2011	2014	No	Yes
Aksum University	Tigray	2006	2009	No	Yes
Ambo University	Oromiya	2009	2012	No	Yes
Arba Minch University	SNNP	2004	2007	No	No
Assosa university	Benishangul Gumuz	2011	2014	No	No
Bahir Dar University	Amhara	1999	2002	No	Yes
Bule Hora University	Oromiya	2011	2014	No	No
Debre Birhan University	Amhara	2006	2009	No	Yes
Debre Markos University	Amhara	2006	2009	No	Yes
Debre Tabor University	Amhara	2011	2014	No	No
Dilla University	Amhara	2006	2009	No	Yes
Dire Dawa University	Dire Dawa	2006	2009	No	Yes
Ethiopian Civil Service College	Addis Ababa	1995	1998	Yes	Yes
FDRE Defense University College	Oromiya	2000	2003	Yes	Yes
Gondar University	Amhara	2003	2006	No	No
Haramaya University	Oromiya	1995	1998	No	No
Hawassa University	SNNP	1999	2002	No	Yes
Jijiga University	Somali	2006	2009	No	No
Jimma University	Oromiya	1999	2002	No	Yes
Kotebe Teachers Education College	Addis Ababa	1979	1982	Yes	Yes
Madawalabu University	Oromiya	2006	2009	No	No
Mekelle University	Tigray	1999	2002	No	Yes
Metu University	Oromiya	2011	2014	No	Yes
Mizan-Tepi University	SNNP	2006	2009	No	No
Samara University	Afar	2007	2010	No	No
Wachemo University	SNNP	2011	2014	No	No
Wolaita Sodo University	SNNP	2006	2009	No	Yes
Woldia university	Amhara	2011	2014	No	No
Wolkite University	SNNP	2011	2014	No	No
Wollega University	Oromiya	2006	2009	No	Yes
Wollo University	Amhara	2006	2009	No	Yes

Notes: The table shows the list of public universities as per the 2013/14 (2006EC) Annual Statistical Abstract from the Ministry of Education, Ethiopia. Accredited Private University is taken from the full list of private universities from the abstracts covering the period 1999/2000 to 2013/14, and including additional campuses of the private universities where possible.

Table 7: Checking for Demographic Changes Due to the Establishment of a Public University, LFS 2013

	(1) Age	(2) Female	(3) Married	(4) Migrant	(5) High School
Exposed District x Exposed Cohort	-0.42	0.027	-0.029	0.0099	0.0065
	(0.30)	(0.024)	(0.024)	(0.028)	(0.015)
Adjusted R-Squared	0.15	0.00034	0.017	0.025	0.014
Observations	24853	24853	24853	24853	24853
Clusters	596	596	596	596	596

Notes: * p<0.1, ** p<0.05, *** p<0.01. Standard errors in brackets. Clustered standard errors are reported, clustered at the district level. Exposed Cohort is an indicator equal to one if the individual is aged 17-20 at the time that the nearest university was established, Exposed District is an indicator equal to one if the respondent is in a district with a public university. The sample is restricted to those aged 17-24 at the time that the nearest university was established. The specifications control for region fixed effects. Figures are rounded to two significant figures.

Table 8: Differential Selection into the Civil Service: ECSS

	(1) Age	(2) Female	(3) Tenure in Service	(4) Undergraduate Degree	(5) Masters Degree
Exposed Cohort x Exposed District	-2.10	0.072	-1.81	-0.26	-0.00031
	(3.53)	(0.13)	(1.38)	(0.36)	(0.0077)
Adjusted R-Squared	0.11	0.022	0.20	0.067	-0.0048
Observations	358	358	358	358	358
Clusters	65	65	65	65	65

Notes: * p<0.1, ** p<0.05, *** p<0.01. Clustered standard errors in brackets. Standard errors are clustered at the district level. The unit of observation is the individual. Exposed Cohort is an indicator equal to one if the civil servant entered the service 0-4 years after the nearest public university was able to graduate students, Exposed District is an indicator equal to one if the respondent is in a district with a public university. The sample is restricted to those that entered the service between 4 years before the nearest public university could start to graduate students and 4 years after. The specifications control for region fixed effects. Figures are rounded to two significant figures.

Table 9: Demographic Characteristics of Districts with a Public University

Variable	(1) No Public University Mean/SE	(2) Public University Mean/SE	T-test Difference (1)-(2)
Population (millions)	0.096	0.121	-0.025
	(0.002)	(0.011)	
Percentage of population female	0.492	0.498	-0.006
	(0.001)	(0.002)	
Percentage of population rural	42.528	22.827	19.701***
	(0.481)	(3.221)	
Percentage of population internal migrants	0.167	0.337	-0.170***
	(0.005)	(0.029)	
Ethnic fractionalisation	0.165	0.263	-0.098***
	(0.007)	(0.039)	
Poverty rate	29.729	25.530	4.199**
	(0.512)	(1.869)	
Remoteness	6.840	2.793	4.047***
	(0.221)	(0.496)	
Has an accredited private university	0.022	0.327	-0.304***
	(0.006)	(0.068)	
N	677	49	
F-test of joint significance (F-stat)			7.354***
F-test, number of observations			726

Notes: * p<0.1, ** p<0.05, *** p<0.01. Standard errors in parentheses. Robust standard errors are reported. The unit of observation is the district. The private university data is gathered from the Ministry of Education Annual Statistical Abstracts. The district-characteristics data is gathered from census records from the 2007 census. Remoteness is the average travel time in hours to the nearest urban centre with at least 50,000 inhabitants. Ethnic fractionalisation represents the probability that two randomly selected individuals from the districts are from different ethnicities (Fearon, 2003). Poverty rate is between 0 and 100 - the percentage of the population living below the national poverty line. T-tests are conducted controlling for region fixed effects. Figures are rounded to two significant figures.

Table 10: Characteristics of Individuals in Districts with and without a Public University, Labor Force Survey 2013

	(1) No Public University		(2) Public University		T-test Difference
Variable	N/[Clusters]	Mean/SE	N/[Clusters]	Mean/SE	(1)-(2)
Age	61951	36.85	26372	35.85	1.00***
	[547]	(0.10)	[49]	(0.31)	
1[Female]	61951	0.51	26372	0.51	-0.00
	[547]	(0.00)	[49]	(0.00)	
1[Married]	61951	0.77	26372	0.66	0.11***
	[547]	(0.00)	[49]	(0.02)	
1[Migrate for work]	61951	0.05	26372	0.16	-0.10***
	[547]	(0.00)	[49]	(0.02)	
1[Migrate for education]	61951	0.01	26372	0.04	-0.03***
	[547]	(0.00)	[49]	(0.01)	
1[High-school education]	61490	0.06	26197	0.15	-0.10***
	[547]	(0.00)	[49]	(0.02)	
1[Post-secondary education]	61490	0.04	26197	0.15	-0.10***
	[547]	(0.00)	[49]	(0.02)	
1[Currently employed]	61951	0.87	26372	0.77	0.11***
	[547]	(0.00)	[49]	(0.02)	
Hours of work, all jobs last week	51650	33.23	18988	38.53	-5.30***
	[547]	(0.39)	[49]	(1.38)	
1[Public sector employee]	51650	0.05	18988	0.15	-0.10***
	[547]	(0.00)	[49]	(0.02)	
F-test of joint significance (F-stat)					3.85***

Notes: * p<0.1, ** p<0.05, *** p<0.01. Standard errors in brackets. Clustered standard errors are reported. Standard errors are clustered at the district level. The unit of observation is the individual. The sample is restricted to those aged between 21 and 65 only. 1[Migrate for work] is equal to one if the individual has migrated for employment or job search. 1[Migrate for education] is equal to one if the individual migrated for education. 1[High-school education] is equal to one if the highest education level of the individual is secondary level. 1[Post-secondary education] is equal to one if the individual has a tertiary education. 1[Currently employed] is equal to one if the respondent is currently employed, relative to unemployed or inactive. Hours of work, all jobs last week is the reported number of hours that the respondent worked across all jobs in the last week. 1[Public sector employee] is equal to one if the individual is working in the public sector, conditional on the individual being employed. T-tests are conducted controlling for region fixed effects. Figures are rounded to two significant figures.

Table 11: Quantile Regression: Log Hourly Wages, Public and Private

	(1) Q=0.2	(2) Q=0.4	(3) Q=0.6	(4) Q=0.8
Exposed District x Exposed Cohort	0.00	-0.33	-0.26***	-0.24***
	(0.0029)	(0.24)	(0.080)	(0.079)
Exposed District x Exposed Cohort x 1[Public]	-0.10	0.33	0.24***	0.18*
	(0.099)	(0.27)	(0.077)	(0.100)
District FE	Yes	Yes	Yes	Yes
Cohort of Birth FE	Yes	Yes	Yes	Yes
Individual controls	Yes	Yes	Yes	Yes
Observations	10136			
Clusters	581			

Notes: * p<0.1, ** p<0.05, *** p<0.01. Quantile regression estimates. Standard errors in parentheses. Standard errors clustered at the district level. The unit of observation is the respondent. Exposed Cohort is an indicator equal to one if the individual is aged 17-20 at the time that the nearest university was established, Exposed District is an indicator equal to one if the respondent is in a district with a public university. The sample is restricted to those aged 17-24 at the time that the nearest university was established. 1Public] is equal to one if the individual works in the public sector. The dependent variable is the log hourly wage. Each column shows the effect for each quintile of log hourly wages. Figures are rounded to two significant figures.

Table 12: Log Hourly Wages, Post-Secondary-Educated and Non-Post-Secondary-Educated

	(1)	(2)	(3)	(4)
	Baseline	Post-Secondary	All Controls	Monthly Earnings USD
Exposed Cohort	-0.027	0.035	0.064	3.79**
	(0.037)	(0.028)	(0.054)	(1.73)
Exposed District x Exposed Cohort	-0.14**	-0.19***	-0.13**	-4.04**
	(0.061)	(0.069)	(0.054)	(1.75)
1[Post-secondary education]		1.92***	1.41***	41.2***
		(0.052)	(0.066)	(1.57)
Exposed District x 1[Post-secondary]		-0.74***	-0.38***	5.70*
		(0.11)	(0.10)	(3.20)
Exposed Cohort x 1[Post-secondary]		-0.20***	-0.14**	-7.06***
		(0.055)	(0.062)	(1.70)
Exposed Cohort x Exposed District x 1[Post-secondary]		0.22**	0.16	2.61
		(0.10)	(0.100)	(3.36)
Cohort of Birth FE	No	No	Yes	Yes
Individual controls	No	No	Yes	Yes
District FE	No	No	Yes	Yes
Adjusted R-Squared	0.12	0.47	0.62	0.54
Observations	10136	10136	10136	10136
Clusters	581	581	581	581

Notes: * p<0.1, ** p<0.05, *** p<0.01. Standard errors in parentheses. Standard errors clustered at the district level. The unit of observation is the respondent. Exposed Cohort is an indicator equal to one if the individual is aged 17-20 at the time that the nearest university was established, Exposed District is an indicator equal to one if the respondent is in a district with a public university. The sample is restricted to those aged 17-24 at the time that the nearest university was established. The dependent variable is the log hourly wage. Column (1) shows the effect of the interaction between Exposed Cohort and Exposed District on the log hourly wage; column (2) interacts the variables with an indicator for whether the individual has a post-secondary-level education; column (3) adds district fixed effects, cohort fixed effects, and an indicator for whether the respondent is a female; column (4) changes the dependent variable to the monthly wage in USD. Figures are rounded to two significant figures.

Table 13: Measuring Management Practices Using the World Management Survey

Aggregate Practice	Topic	Question	Score 1	Score 3	Score 5
Screening	**Flexibility**	Does your Directorate make efforts to adjust to the specific needs and specific requirements of communities, clients, or other stakeholders?	The directorate uses the same procedures no matter what.	The directorate tailors procedures to the specific needs of its stakeholders, but struggles when those needs are complex.	The directorate tailors all procedures to the specific needs of its stakeholders. The evolution of those needs results in adaptation to plans, project and policies.
Screening	**Flexibility**	How flexible would you say your Directorate is in terms of responding to new and improved work practices or reforms?	New practices are not adopted/ integrated in the directorate.	New ideas or practices are adopted, but in an informal and/ or isolated manner. The directorate encourages the adoption of new practices, however it is slow to integrate them into its operations (more than a year).	The adoption of new ideas and practices is an integral part of the directorate?s work. New practices are regularly reviewed and considered, and once adopted and integrated across the directorate within 6 months.
Screening	**Staffing**	Do you think the management of your Directorate think about attracting talented people to your Directorate and then doing their best to keep them? For example, by ensuring they are happy and engaged with their work.	Directorate does not put emphasis on talent	Senior management believes that attracting and developing talent is important, but there is no clear system for identifying, attracting or retaining such talent.	Senior management believes that attracting and developing talent is important. There is a clear system for identifying and attracting talent, developing and retaining talent.
Screening	**Staffing**	If two senior level staff joined your Directorate five years ago and one was much better at their work than the other, would he/she be promoted through the service faster?	No promotion system (no one in the organisation has been promoted for years) The promotion system is based on tenure	The promotion system is based on performance. Organisation may have internal limitations (e.g. few position openings), but do everything to get around them (e.g. extra training).	Promotion system is based on performance. Organisation actively identifies, develops and promotes top performers. Regular assessments, clear set of indicators and personalised career plans for individuals (regularly revised).
Other	**Roles**	When staff in your Directorate are given tasks in their daily work, how much discretion do they have to carry out their assignments? Can you give me an example?	How officers carry out their assignments is decided by senior managers. Officers have no say.	How officers carry out their assignments is jointly decided by the officer and senior managers. Senior managers tend to drive the decisions.	Officers have complete autonomy in deciding how to carry out their tasks.
Other	**Roles**	Can most staff in your Directorate make substantive contributions to the policy formulation and implementation process?	Staff do not contribute to policy formulation, nor to decisions about implementation.	Staff can contribute to policy formulation and decisions about implementation, but there is no formal forum through which to do this. Contributions typically only occur when problems arise.	Management expects all staff to contribute to policy formulation and decisions about implementation (formally or informally), and considers this part of their duties.
Other	**Roles**	Is the workload of achieving your Directorate?s targets evenly distributed across its different employees, or do some groups consistently shoulder a greater burden than others?	A small minority of staff undertake the vast majority of work within the directorate.	The burden of the directorate?s work is more or less distributed equally among staff. A small minority get away with working significantly less than others.	The burden of the directorate?s work is distributed equally among staff. Tasks are assigned in such a way that the amount of time required and the level of difficulty are balanced out so no member of staff finds him/herself overburdened.
Other	**Roles**	Thinking about all the projects that your Directorate has been involved in since your appointment here, would you say that managers and supervisors try to use the right staff for the right job?	Staff are allocated to tasks randomly.	Managers try to use the right staff for the right job but do not go to great lengths to ensure this, or are met with institutional constraints which may prevent them from doing so.	The right staff are always used for a task. Allocation of tasks is based on staffs? documented skills and competencies.
Other	**Staff involvement**	How do problems in your directorate get exposed and fixed?	Ad-hoc, no set process for improvement Deal with problems as they arise without following an established procedureOnce fixed, no further action taken No suggestions from staff	Existing process to deal with problems Improvements made through meetings Focus on finding solutions, not prevention of future problems Suggestions from staff involved through meetings (formal or informal)	Exposing problems and suggesting solutions and improvements is part of all staffs? daily duty. Continuous improvement is part of the culture of the organisation.
Other	**Staff involvement**	What kind of feedback do you get in staff meetings?	No feedback from staff.	Staff provide feedback in meetings but in an unstructured manner. Focus on bad performance.	Staff provide the feedback on which action plans will be based. Focus on both good and bad performance. Details of the meetings are recorded and communicated to all staff.
Other	**Staff involvement**	Let?s say you?ve agreed to a follow up plan at one of your meetings, what would happen if the plan wasn?t enacted?	No action taken. No changes made in the operations process.	Failure can be found in regular meetings (weekly, even monthly for long-term plans) or at standard points before the deadline. Plans can be altered in order to achieve expected results on time.	In addition to 4, tools can be checked up and reported to the manager in charge. Meetings (formal/ informal) are held to look into the root causes of problems and preventive actions are taken for future similar task.

Table 14: Measuring Management Practices Using the World Management Survey

Aggregate Practice	WMS Topic	Question	Score 1	Score 3	Score 5
Incentives	**Targeting**	Does your Directorate have a clear set of targets derived from the organization?s goals and objectives? Are they used to determine your work schedule?	The directorate does not have defined targets.	Targets are assigned to the directorate, as well as to the manager and employee levels, and these are generally well understood by mid-level staff. However the tasks assigned to staff are not always related to those targets.	Targets are clearly defined for the directorate, manager, and employee levels, and are well understood by all staff. All tasks are directly derived from the targets, which are regularly reviewed to ensure they remain on track.
Incentives	**Targeting**	When you arrive at work each day, do you and your colleagues know what their individual roles and responsibilities are in achieving the organisation?s goals?	Staff do not know what their roles and responsibilities are.	Staff have a good idea of their roles and responsibilities but it is not always clear how they contribute to their organisation?s goals.	Staff have a very good understanding of their roles and responsibilities. Their own roles and goals are clearly interconnected to those of their organisation.
Incentives	**Targeting**	How are targets and performance measures communicated to staff in your directorate?	Neither targets nor performance measures are communicated to staff.	Targets and performance measures are formally communicated to managers and team leaders.	Targets and performance measures are formally communicated and understood by all staff.
Incentives	**Monitoring**	In what kind of ways does your Directorate track how well it is delivering services? Can you give me an example?	Directorate does not track performance.	Directorate tracks a number of performance indicators. These are seen and reviewed by senior management only.	Full set of indicators are tracked formally and continuously. Reviews are conducted regularly and involve representative of all directorate staff groups. The results of the review are formally communicate to all directorate staff.
Incentives	**Monitoring**	Are you involved in performance review for your Directorate? If so, how often does this occur?	Not involved in performance review;	Bi-annually	Monthly
Incentives	**Performance incentives**	How would underperformance be tolerated in your Directorate? Can you give me an example of how such a case would be dealt with?	Poor performers stay in their positions (no consequences).	Poor performance Is identified through evaluation and is addressed through concrete action. Although this applies to most staff, some individuals/staff groups get away with it.	Poor performers are identified through regular reviews and are put on a formal performance improvement plan immediately. This applies to all staff.
Incentives	**Performance incentives**	Given past experience, have members of [respondent?s organisation] been disciplined for breaking the rules of the civil service?	There are no consequences for bad behaviour/ breaking the rules.	Bad behaviour is addressed through concrete action, but the underlying issues are not addressed.	Bad behaviour/ breaking the rules is addressed through concrete action. If any employee breaks the rules, the underlying issues will be identified and rectified. This applies to all employees.
Incentives	**Performance incentives**	Does your Directorate use performance, targets, or indicators for tracking and rewarding (financially or non-financially) the performance of its employees?	Staff are rewarded equally (or not rewarded) irrespective of performance. Individual performance is not tracked formally	There is a formal staff evaluation system in place and performance is rewarded (financially or non-financially). However, there are no clear system or criteria for rewarding staff	There is a formal staff evaluation system and performance is rewarded (financially or non- financially). Rewards are given as a consequence of well-defined and monitored individual achievements. This applied to all staff.

Table 15: Robustness: Main Regressions with Finer Cohort Ranges (20 v 21)

	(1) Post-Secondary Education	(2) Employed in Last 12 Months High-Skilled	(3) Log Hourly Wage Public-Private	(4) Log Hourly Wage Education
Exposed Cohort	-0.014	0.15***	0.67**	0.64**
	(0.026)	(0.043)	(0.28)	(0.26)
Exposed District x Exposed Cohort	0.053**	-0.080	-0.27***	-0.22*
	(0.021)	(0.050)	(0.097)	(0.11)
1[Public Sector]			1.96***	
			(0.17)	
Exposed District x 1[Public Sector]			-1.07***	
			(0.22)	
Exposed Cohort x 1[Public Sector]			0.020	
			(0.094)	
Exposed District x Exposed Cohort x 1[Public Sector]			0.49***	
			(0.15)	
1[Post-secondary education]				1.17***
				(0.16)
Exposed District x 1[Post-secondary]				-0.65***
				(0.19)
Exposed Cohort x 1[Post-secondary]				-0.11
				(0.15)
Exposed District x Exposed Cohort x 1[Post-secondary]				0.56***
				(0.20)
Cohort of Birth FE	Yes	Yes	Yes	Yes
Individual controls	Yes	Yes	Yes	Yes
District FE	Yes	Yes	Yes	Yes
Adjusted R-Squared	0.089	0.21	0.74	0.67
Observations	7220	1179	2997	2980
Clusters	586	215	526	525

Notes: * p<0.1, ** p<0.05, *** p<0.01. Standard errors in parentheses. Standard errors clustered at the district level. The unit of observation is the respondent. Exposed Cohort is an indicator equal to one if the individual is aged 20 at the time that the nearest university was established. Exposed District is an indicator equal to one if the respondent is in a district with a public university. The sample is restricted to those aged between 20 and 21 at the time that the nearest university was established. The dependent variable in column (1) is an indicator for whether the individual has a post-secondary-level education; in column (2) an indicator for whether the individual has been employed within the last 12 months; in columns (3) and (4) the log wage-per-hour. In columns (1) and (2), individual controls include an indicator for whether the respondent is female. In columns (3) and (4), individual controls also include an indicator for whether the respondent has no formal education, primary level education or secondary-and-above education. All columns include cohort-of-birth fixed effects and district fixed effects. Figures are rounded to two significant figures.

Table 16: Robustness: Cohorts 17-20 v 25-28 at the Time of Establishment

	(1) Post-Secondary Education	(2) Employed in Last 12 Months High-Skilled	(3) Log Hourly Wage Public-Private	(4) Log Hourly Wage Education
Exposed Cohort	-0.030	-0.20	-0.021	0.085
	(0.030)	(0.18)	(0.20)	(0.24)
Exposed District x Exposed Cohort	0.037***	-0.10***	-0.15*	-0.11*
	(0.011)	(0.027)	(0.076)	(0.065)
1[Public Sector]			1.85***	
			(0.15)	
Exposed District x 1[Public Sector]			-0.87***	
			(0.17)	
Exposed Cohort x 1[Public Sector]			0.11	
			(0.078)	
Exposed District x Exposed Cohort x 1[Public Sector]			0.11	
			(0.12)	
1[Post-secondary education]				1.22***
				(0.13)
Exposed District x 1[Post-secondary]				-0.42***
				(0.16)
Exposed Cohort x 1[Post-secondary]				-0.21*
				(0.12)
Exposed District x Exposed Cohort x 1[Post-secondary]				0.20
				(0.15)
Cohort of Birth FE	Yes	Yes	Yes	Yes
Individual controls	Yes	Yes	Yes	Yes
District FE	Yes	Yes	Yes	Yes
Adjusted R-Squared	0.11	0.16	0.75	0.69
Observations	22799	3278	9022	8972
Clusters	596	337	576	576

Notes: * p<0.1, ** p<0.05, *** p<0.01. Standard errors in parentheses. Standard errors clustered at the district level. The unit of observation is the respondent. Exposed Cohort is an indicator equal to one if the individual is aged 17-20 at the time that the nearest university was established. Exposed District is an indicator equal to one if the respondent is in a district with a public university. The sample is restricted to those aged between 17-20 and those aged between 25-28 at the time that the nearest university was established. The dependent variable in column (1) is an indicator for whether the individual has a post-secondary-level education; in column (2) an indicator for whether the individual has been employed within the last 12 months; in columns (3) and (4) the log wage-per-hour. In columns (1) and (2), individual controls include an indicator for whether the respondent is female. In columns (3) and (4), individual controls also include an indicator for whether the respondent has no formal education, primary level education or secondary-and-above education. All columns include cohort-of-birth fixed effects and district fixed effects. Figures are rounded to two significant figures.

Table 17: Robustness: District Exposure by Remoteness

	(1) Post-Secondary Education	(2) Employed in Last 12 Months High-Skilled	(3) Log Hourly Wage Public-Private	(4) Log Hourly Wage Education
Exposed Cohort	-0.024	0.0015	0.0033	0.030
	(0.015)	(0.032)	(0.059)	(0.086)
Exposed District x Exposed Cohort	0.028***	-0.028	-0.13**	-0.083
	(0.0092)	(0.028)	(0.053)	(0.061)
1[Public Sector]			2.20***	
			(0.12)	
Exposed District x 1[Public Sector]			-1.07***	
			(0.13)	
Exposed Cohort x 1[Public Sector]			0.057	
			(0.065)	
Exposed District x Exposed Cohort x 1[Public Sector]			0.12	
			(0.085)	
1[Post-secondary education]				1.41***
				(0.11)
Exposed District x 1[Post-secondary]				-0.64***
				(0.13)
Exposed Cohort x 1[Post-secondary]				-0.12
				(0.13)
Exposed District x Exposed Cohort x 1[Post-secondary]				0.14
				(0.14)
Cohort of Birth FE	Yes	Yes	Yes	Yes
Individual controls	Yes	Yes	Yes	Yes
District FE	Yes	Yes	Yes	Yes
Adjusted R-Squared	0.11	0.15	0.75	0.68
Observations	24853	4015	10192	10136
Clusters	596	353	581	581

Notes: * p<0.1, ** p<0.05, *** p<0.01. Standard errors in parentheses. Standard errors clustered at the district level. The unit of observation is the respondent. Exposed Cohort is an indicator equal to one if the individual is aged 17-20 at the time that the nearest university was established. Exposed District is an indicator equal to one if the respondent is in a district that is less than 3.1 hours away from the nearest large urban area (the median travel time). The sample is restricted to those aged between 17-24 at the time that the nearest university was established. The dependent variable in column (1) is an indicator for whether the individual has a post-secondary-level education; in column (2) an indicator for whether the individual has been employed within the last 12 months; in columns (3) and (4) the log wage-per-hour. In columns (1) and (2), individual controls include an indicator for whether the respondent is female. In columns (3) and (4), individual controls also include an indicator for whether the respondent has no formal education, primary level education or secondary-and-above education. All columns include cohort-of-birth fixed effects and district fixed effects. Figures are rounded to two significant figures.

Table 18: Robustness: District Exposure by Distance

	(1) Post-Secondary Education	(2) Employed in Last 12 Months High-Skilled	(3) Log Hourly Wage Public-Private	(4) Log Hourly Wage Education
Exposed Cohort	-0.020	0.015	-0.049	0.010
	(0.015)	(0.031)	(0.059)	(0.087)
Exposed District x Exposed Cohort	0.019**	-0.050*	-0.025	-0.028
	(0.0093)	(0.027)	(0.055)	(0.061)
1[Public Sector]			2.04***	
			(0.12)	
Exposed District x 1[Public Sector]			-0.78***	
			(0.14)	
Exposed Cohort x 1[Public Sector]			0.13**	
			(0.063)	
Exposed District x Exposed Cohort x 1[Public Sector]			-0.065	
			(0.089)	
1[Post-secondary education]				1.29***
				(0.093)
Exposed District x 1[Post-secondary]				-0.45***
				(0.12)
Exposed Cohort x 1[Post-secondary]				-0.16
				(0.11)
Exposed District x Exposed Cohort x 1[Post-secondary]				0.16
				(0.13)
Cohort of Birth FE	Yes	Yes	Yes	Yes
Individual controls	Yes	Yes	Yes	Yes
District FE	Yes	Yes	Yes	Yes
Adjusted R-Squared	0.11	0.15	0.74	0.68
Observations	24853	4015	10192	10136
Clusters	596	353	581	581

Notes: * p<0.1, ** p<0.05, *** p<0.01. Standard errors in parentheses. Standard errors clustered at the district level. The unit of observation is the respondent. Exposed Cohort is an indicator equal to one if the individual is aged 17-20 at the time that the nearest university was established. Exposed District is an indicator equal to one if the respondent is in a district that is less than 40 kilometres away from the nearest public university (the median distance). The sample is restricted to those aged between 17-24 at the time that the nearest university was established. The dependent variable in column (1) is an indicator for whether the individual has a post-secondary-level education; in column (2) an indicator for whether the individual has been employed within the last 12 months; in columns (3) and (4) the log wage-per-hour. In columns (1) and (2), individual controls include an indicator for whether the respondent is female. In columns (3) and (4), individual controls also include an indicator for whether the respondent has no formal education, primary level education or secondary-and-above education. All columns include cohort-of-birth fixed effects and district fixed effects. Figures are rounded to two significant figures.

Table 19: Robustness: District Exposure by Remoteness and Distance

	(1) Post-Secondary Education	(2) Employed in Last 12 Months High-Skilled	(3) Log Hourly Wage Public-Private	(4) Log Hourly Wage Education
Exposed Cohort	-0.023	0.0076	-0.027	0.026
	(0.015)	(0.030)	(0.059)	(0.083)
Exposed District x Exposed Cohort	0.029***	-0.043*	-0.078	-0.069
	(0.0095)	(0.025)	(0.060)	(0.064)
1[Public Sector]			2.10***	
			(0.12)	
Exposed District x 1[Public Sector]			-1.01***	
			(0.14)	
Exposed Cohort x 1[Public Sector]			0.090	
			(0.059)	
Exposed District x Exposed Cohort x 1[Public Sector]			0.061	
			(0.085)	
1[Post-secondary education]				1.35***
				(0.097)
Exposed District x 1[Post-secondary]				-0.61***
				(0.12)
Exposed Cohort x 1[Post-secondary]				-0.18*
				(0.10)
Exposed District x Exposed Cohort x 1[Post-secondary]				0.24*
				(0.13)
Cohort of Birth FE	Yes	Yes	Yes	Yes
Individual controls	Yes	Yes	Yes	Yes
District FE	Yes	Yes	Yes	Yes
Adjusted R-Squared	0.11	0.15	0.75	0.68
Observations	24853	4015	10192	10136
Clusters	596	353	581	581

Notes: * p<0.1, ** p<0.05, *** p<0.01. Standard errors in parentheses. Standard errors clustered at the district level. The unit of observation is the respondent. Exposed Cohort is an indicator equal to one if the individual is aged 17-20 at the time that the nearest university was established. Exposed District is an indicator equal to one if the respondent is in a district that is less than 40 kilometres away from the nearest public university (the median distance) and less than 3.1 hours away from the nearest large urban area (the median travel time). The sample is restricted to those aged between 17-24 at the time that the nearest university was established. The dependent variable in column (1) is an indicator for whether the individual has a post-secondary-level education; in column (2) an indicator for whether the individual has been employed within the last 12 months; in columns (3) and (4) the log wage-per-hour. In columns (1) and (2), individual controls include an indicator for whether the respondent is female. In columns (3) and (4), individual controls also include an indicator for whether the respondent has no formal education, primary level education or secondary-and-above education. All columns include cohort-of-birth fixed effects and district fixed effects. Figures are rounded to two significant figures.

Table 20: World Values Survey: Intrinsic Motivation

	Measure 1			Measure 2		
	(1)	(2)	(3)	(4)	(5)	(6)
	Unconditional	Conditional	Public and Private	Unconditional	Conditional	Public and Private
Regional Exposure	0.0013	0.0011	0.000078	0.041	0.039	0.031
	(0.012)	(0.012)	(0.013)	(0.025)	(0.025)	(0.025)
1[Public]			0.055			0.086
			(0.036)			(0.064)
Regional Exposure x 1[Public]			0.0038			0.11**
			(0.023)			(0.047)
Cohort of Birth FE	Yes	Yes	Yes	Yes	Yes	Yes
Region FE	Yes	Yes	Yes	Yes	Yes	Yes
Individual controls	No	Yes	Yes	No	Yes	Yes
Adjusted R-Squared	-0.017	-0.014	-0.012	0.050	0.065	0.075
Observations	1198	1198	1198	1198	1198	1198

Notes: * p<0.1, ** p<0.05, *** p<0.01. OLS estimates. Standard errors in parentheses. Robust standard errors reported. Regional Exposure is the number of universities available in the region when aged 17-20 years old. 1[Public] is an indicator equal to one if the individual works in the public sector. The dependent variable in columns (1) to (3) is equal to one for those that respond 'doing an important job' to "Regardless of whether you're actually looking for a job, which one would you, personally, place first if you were looking for a job?". The dependent variable in columns (4) to (6) is equal to 1 for those that respond 'Very much like me' or 'Like me' to "Using this card, would you please indicate for each description whether that person is very much like you, like you, somewhat like you, not like you, or not at all like you? It is important to this person to help the people nearby; to care for their well-being.". Both are based on Cowley and Smith [2014]. All columns include region fixed effects and cohort fixed effects. Columns (2) to (3) and (5) to (6) also include an indicator for whether the individual is female. Columns (3) and (6) include an indicator for whether the individual works in the public sector. Figures are rounded to two significant figures.

Table 21: Household Wealth Between Cohorts (LSMS)

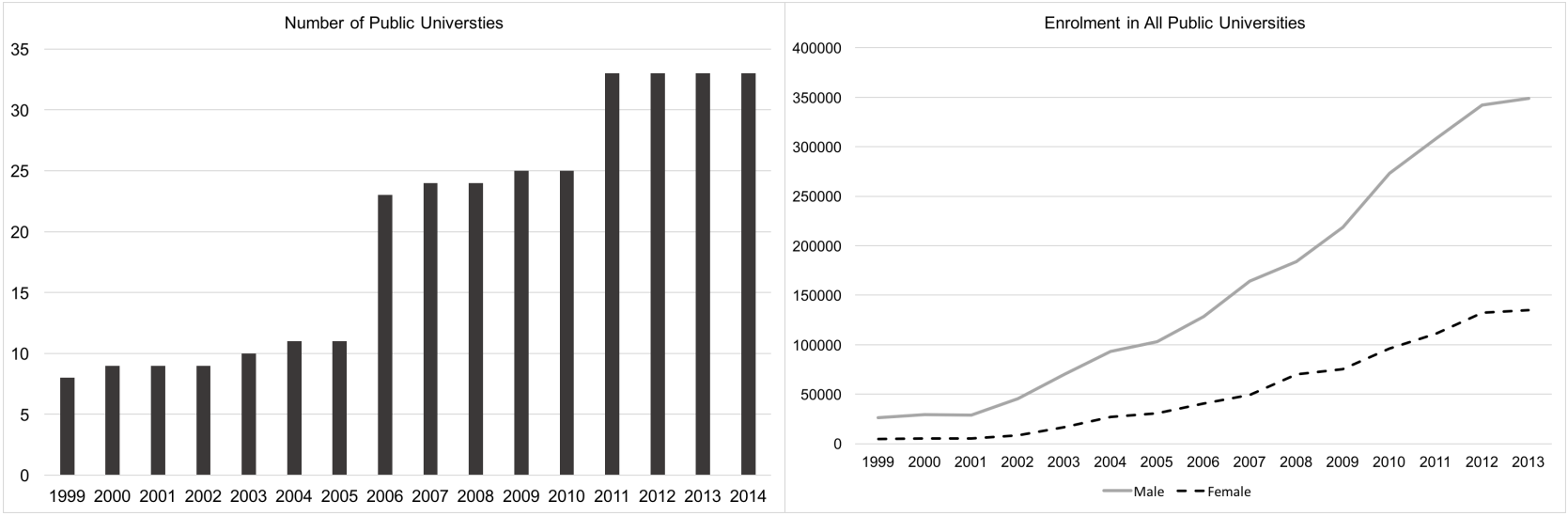
	All Individuals					Post-Secondary-Educated Only				
	(1) 1[Mother Completed Primary Education]	(2) 1[Mother Completed Secondary Education]	(3) 1[Father Completed Primary Education]	(4) 1[Father Completed Secondary Education]	(5) Household Total Listed Assets	(6) 1[Mother Completed Primary Education]	(7) 1[Mother Completed Secondary Education]	(8) 1[Father Completed Primary Education]	(9) 1[Father Completed Secondary Education]	(10) Household Total Listed Assets
Exposed Cohort x Exposed District	0.014	-0.00066	0.019	-0.0091	0.027	0.036	-0.0020	-0.092	-0.052	-2.81
	(0.031)	(0.017)	(0.036)	(0.025)	(1.78)	(0.083)	(0.051)	(0.12)	(0.083)	(4.68)
Cohort FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Individual controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
District FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Adjusted R-Squared	-0.039	-0.058	0.061	0.0028	0.34	-0.086	0.022	0.065	0.0075	0.31
Observations	2298	2298	2299	2299	2358	438	438	439	439	439
Clusters	294	294	294	294	294	136	136	136	136	136

Notes: * p<0.1, ** p<0.05, *** p<0.01. OLS estimates. Clustered standard errors in parentheses. Standard errors clustered at district level. The sample is restricted to those aged between 21 and 65 and to those who were aged between 17 and 24 at the time that their nearest university was established. Exposed Cohort x Exposed District is the interaction between an indicator equal to one if the individual was aged 17-20 at the time that the nearest university was established and an indicator equal to one if the individual lives in a district with a public university. The dependent variable in columns (1) and (6) is an indicator for whether the mother has completed at least primary education; in columns (2) and (7) an indicator for whether the mother has at least completed secondary-level education; in columns (3) and (8) an indicator for whether the father has at least a primary-level education; in columns (4) and (9) an indicator for whether the father has at least a secondary-level education; in columns (5) and (10) the total number of listed assets that the household reports in the LSMS. Columns (1) to (5) display the results for all individuals, while columns (6) to (10) restrict the sample to only those with a post-secondary education. Individual controls include an indicator for whether the respondent is female. All columns control for district fixed effects and cohort-of-birth fixed effects. Figures are rounded to two significant figures.

Table 22: Consumption Effects: LSMS

	(1) Household Monthly Expenditure All	(2) Employed Last 12 Months Inc. Casual	(3) Not Employed Last 12 Months	(4) Post-Secondary Educated Only	(5) Post-Secondary Educated Employed	(6) Post-Secondary Educated Not Employed
Exposed Cohort x Exposed District	-1.09	-6.54	-10.1	-40.1***	-53.0	-41.6
	(10.4)	(23.1)	(15.0)	(15.2)	(33.5)	(27.5)
Cohort of Birth FE	Yes	Yes	Yes	Yes	Yes	Yes
Individual controls	Yes	Yes	Yes	Yes	Yes	Yes
District FE	Yes	Yes	Yes	Yes	Yes	Yes
Adjusted R-Squared	0.29	0.32	0.29	0.36	0.27	0.34
Observations	2348	336	2012	438	195	243
Clusters	294	114	292	136	89	96

Notes: * p<0.1, ** p<0.05, *** p<0.01. OLS estimates. Clustered standard errors in parentheses. Standard errors clustered at district level. The sample is restricted to those aged between 21 and 65 and to those who were aged between 17 and 24 at the time that their nearest university was established. Exposed Cohort x Exposed District is the interaction between an indicator equal to one if the individual was aged 17-20 at the time that the nearest university was established and an indicator equal to one if the individual lives in a district with a public university. The dependent variable in all columns is the monthly household expenditure reported in the LSMS. Column (1) shows the results for all individuals in the sample, column (2) only for those employed over the last 12 months, column (3) for those not employed over the last 12 month, column (4) for all post-secondary-educated individuals, column (5) for post-secondary-educated individuals employed over the last 12 months, and column (6) for post-secondary-educated individuals not employed over the last 12 months. Individual controls include an indicator for whether the respondent is female. All columns control for district fixed effects and cohort-of-birth fixed effects. Figures are rounded to two significant figures.



(a) All Public Universities **(b)** Enrolment in Public Universities

Notes: The left-hand figure shows the total number of public universities from the period 1999 to 2014 academic years according to the Ministry of Education Annual Statistical Abstracts, Ethiopia. The right-hand figure shows the enrolment in public universities by gender for the period 1999 to 2013 academic years; the black dashed line shows female enrolment and the solid grey line shows male enrolment.

Figure 1: Number of Public Universities and Enrolment, 1999-2013

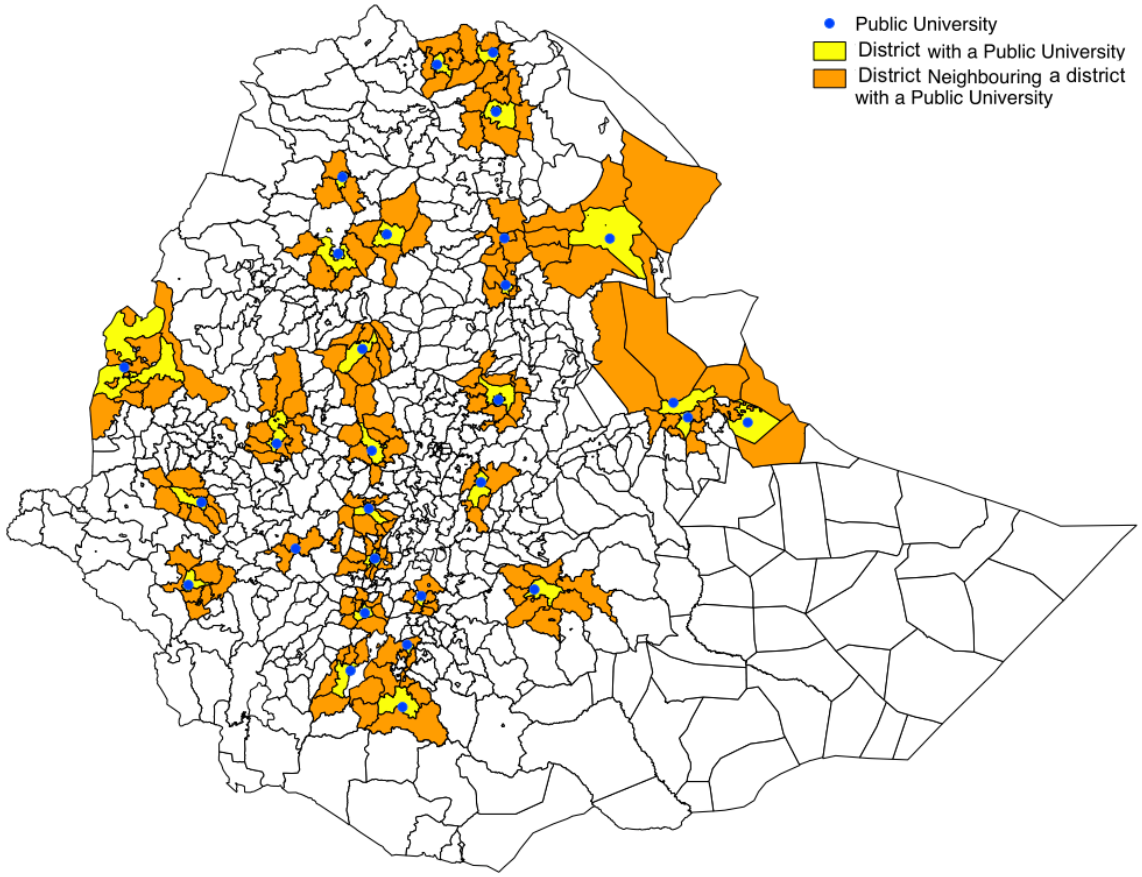
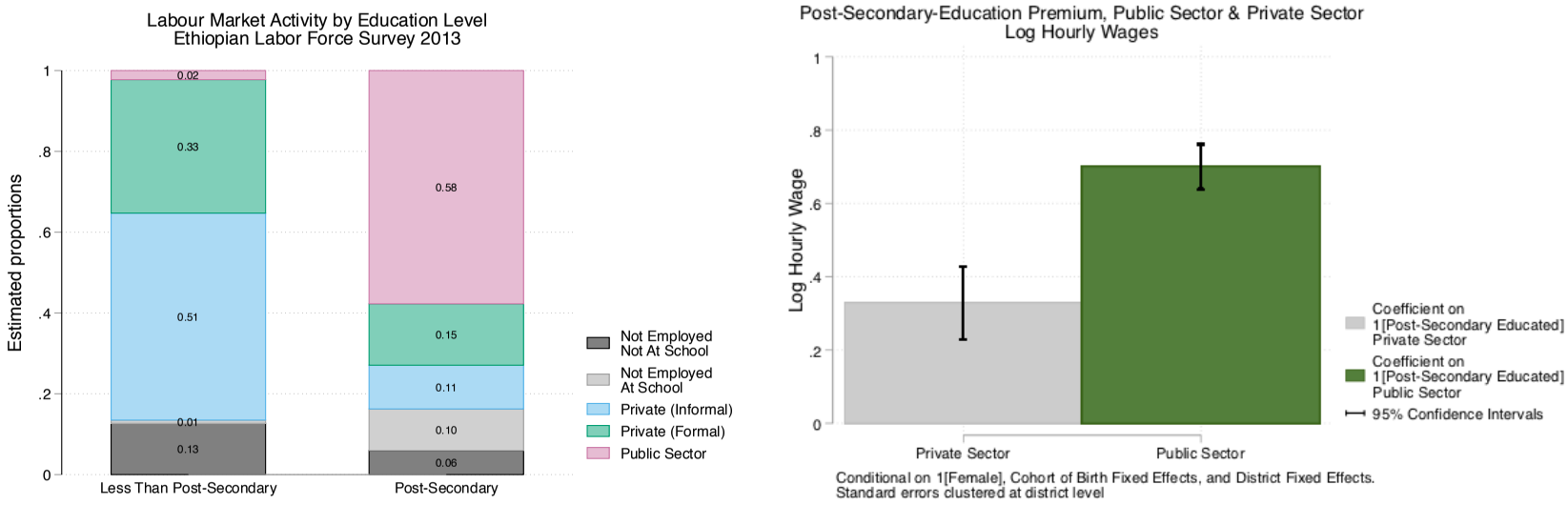


Figure 2: Public Universities, Districts and Neighbouring Districts, Excluding Addis Ababa



(a) Labour Market Activity, by Skill

(b) Post-Secondary Wage Premium, Public and Private

Notes: The left figure shows the labour-market activity of working-age individuals in the Labor Force Survey 2013, by whether they have a post-secondary education or not. The proportions are calculated using weights based on the inverse of the Labour Force Survey 2013 primary sampling unit population. The right figure shows the coefficient on an indicator for whether the individual has a post-secondary education or not in a regression of log hourly wages on education, an indicator for whether the individual is female, cohort-of-birth fixed effects, and district fixed effects. The regression is conducted separately for the private sector (left bar, grey) and public sector (right bar, green). The black lines are the 95% confidence intervals, using standard errors clustered at the district level.

Figure 3: Labour Market Activity by Education Level and Education Wage Premium

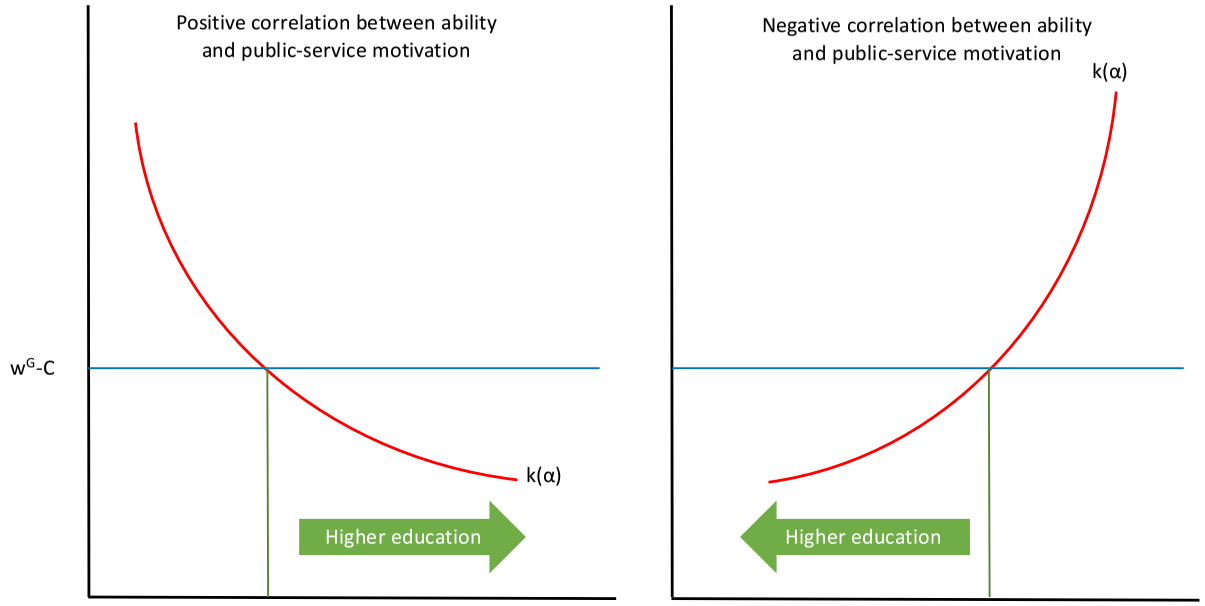


Figure 4: The Distribution of Ability in Higher Education with Positive (left) and Negative (right) Correlation Between Ability and Public-Service Motivation

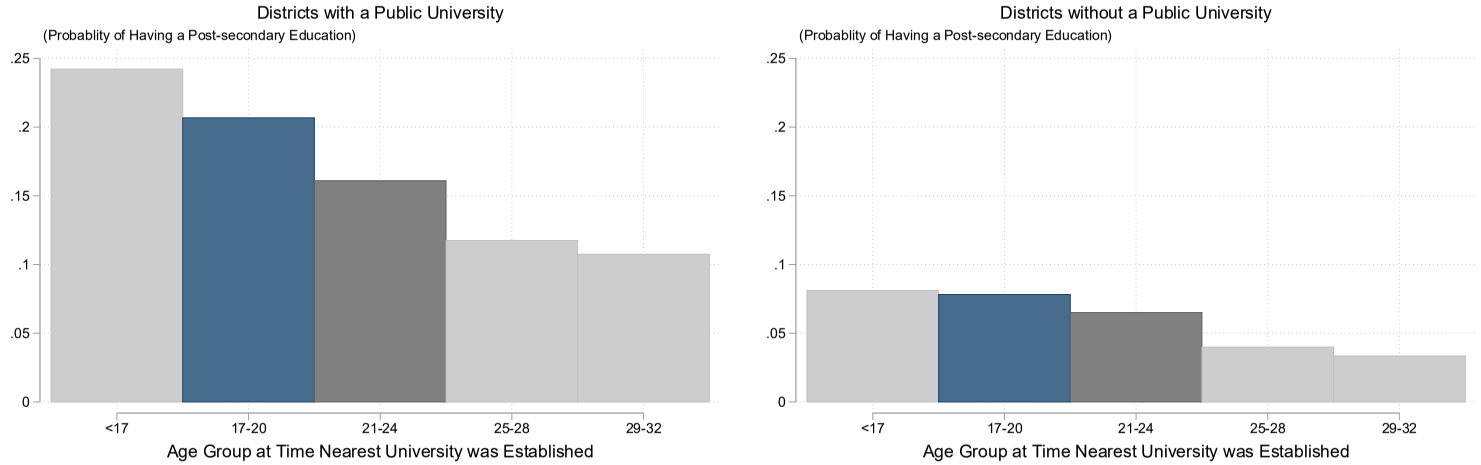
LFS Identification Strategy

		District Has **Public University** Within Borders (Exposed District)	District Does Not Have **Public University** Within Borders (Unexposed District)
		Exposed District	**Unexposed District**
Aged **17-20** When Nearest University Established	**Exposed Cohort**	(1)	(2)
Aged **21-24** When Nearest University Established	**Unexposed Cohort**	(3)	(4)

ECSS Identification Strategy

		District Has **Public University** Within Borders	District Does Not Have **Public University** Within Borders
		Exposed District	**Unexposed District**
Entered the Civil Service **0-4 Years After** Nearest University Produces First Graduates (Exposed Cohort)	**Exposed Cohort**	(5)	(6)
Entered the Civil Service **0-4 Years Before** Nearest University Produces First Graduates (Exposed Cohort)	**Unexposed Cohort**	(7)	(8)

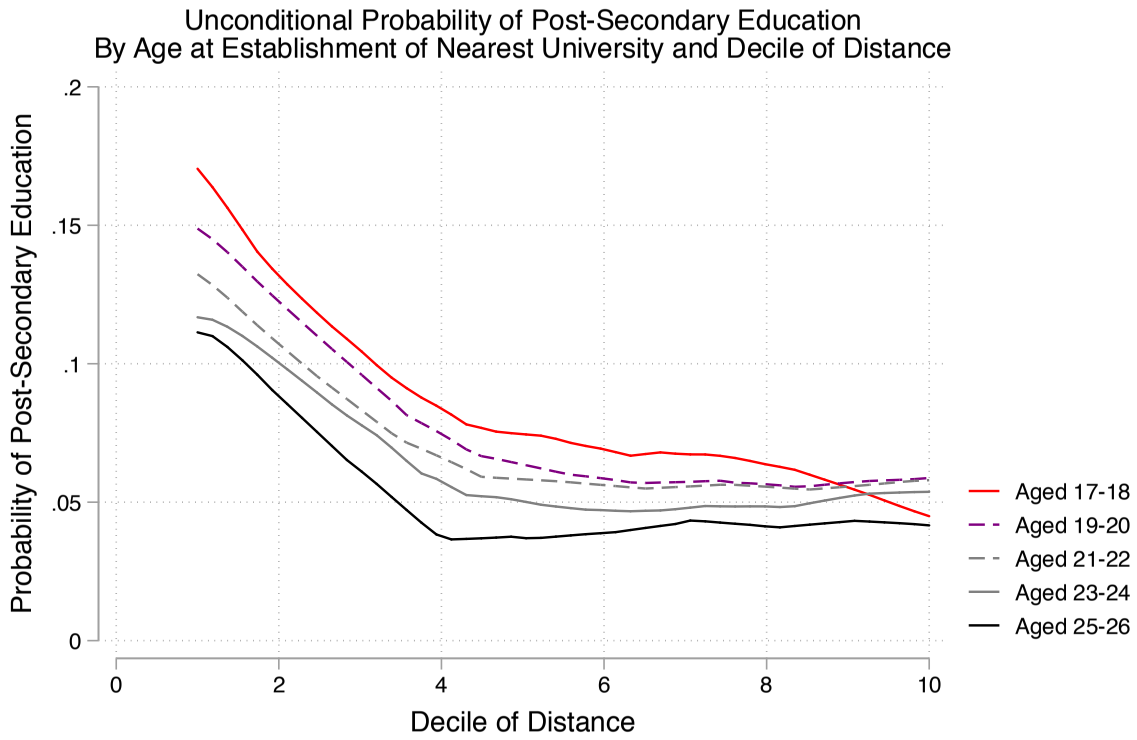
Figure 5: Illustration of Identification Strategy



(a) Exposed Districts **(b)** Unexposed Districts

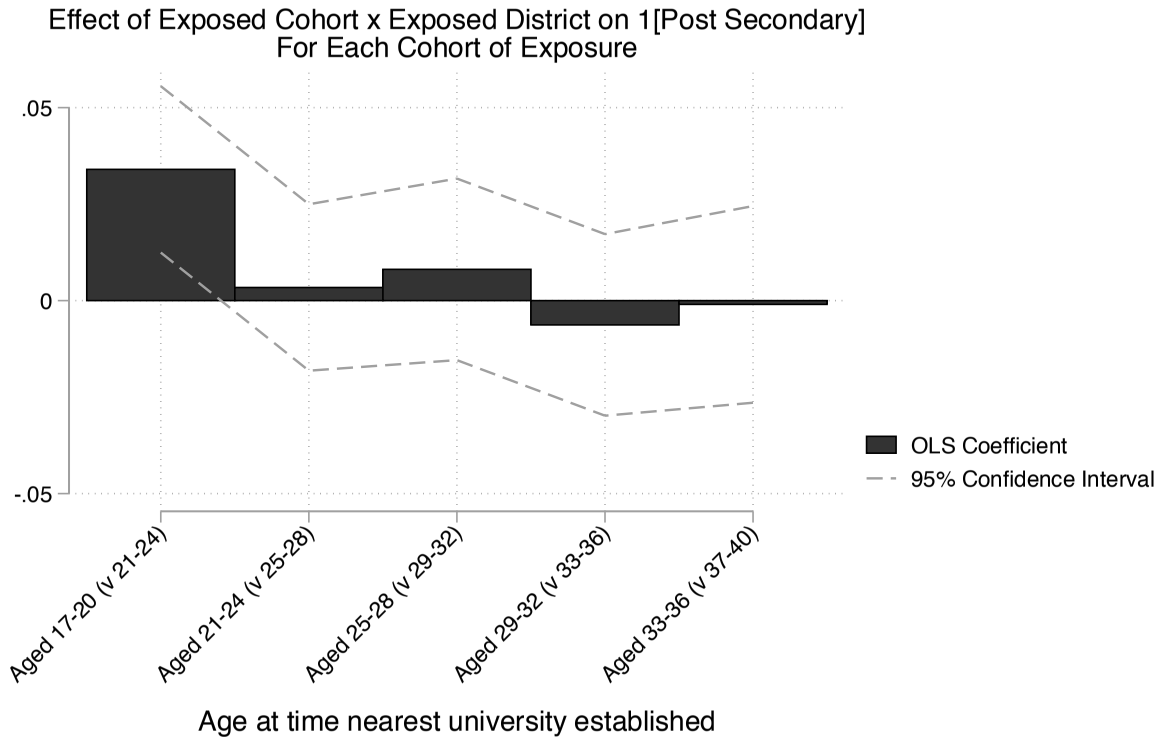
Notes: The figure shows the unconditional probability of having an undergraduate degree by the age of the respondent at the time that the nearest university was fully established. The probabilities are calculated using weights based on the inverse of the Labour Force Survey 2013 primary sampling unit population. The left figure show those districts that have a public university. The right figure shows districts that do not have a public university.

Figure 6: Education Attainment by Cohort and District Exposure



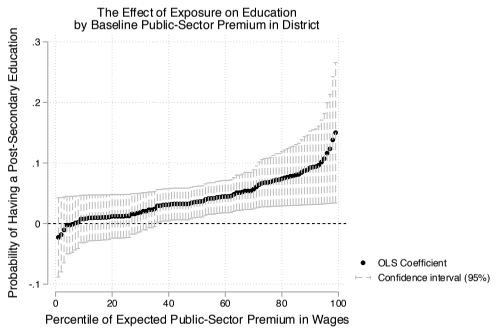
Notes: The figure shows the unconditional probability of having a post-secondary education for each decile of distance to the nearest university and for each age group at the time that the nearest university was established. The probabilities are connected using a kernel-weighted local polynomial, using an Epanechnikov kernel function and a polynomial of degree 1.

Figure 7: Education Attainment by Cohort and District Exposure (Continuous Variation)

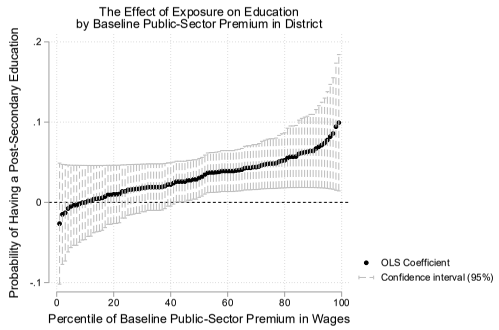


Notes: The figure shows the OLS coefficient (black bars) and 95% confidence intervals (grey dashed lines) for $1[University_d] \times 1[Exposed_{i,c,d}]$ in a regression of an indicator for whether the respondent has a post-secondary education on $1[University_d]$, $1[Exposed_{i,c,d}]$, and $1[University_d] \times 1[Exposed_{i,c,d}]$. The regressions also condition on district fixed effects, cohort-of-birth fixed effects, and an indicator for whether the respondent is female. The figure shows this for each cohort of exposure, relative to the previous cohort, up to aged 40 at the time that the nearest university was established.

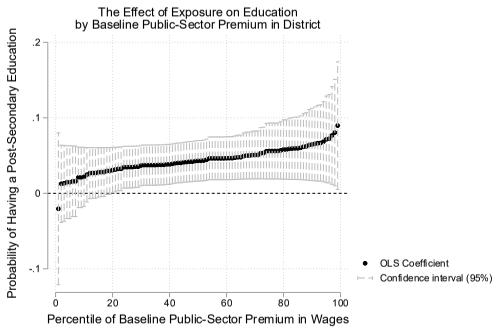
Figure 8: Control Experiment for Each Cohort of Exposure



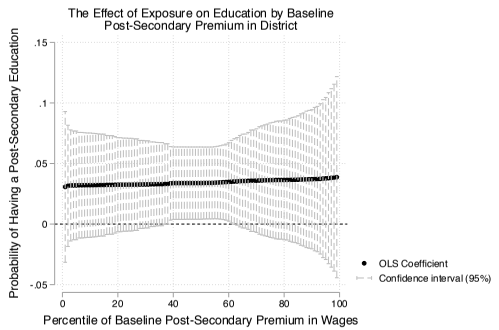
(a) Expected Public Wage Premium



(b) Raw Public Wage Premium



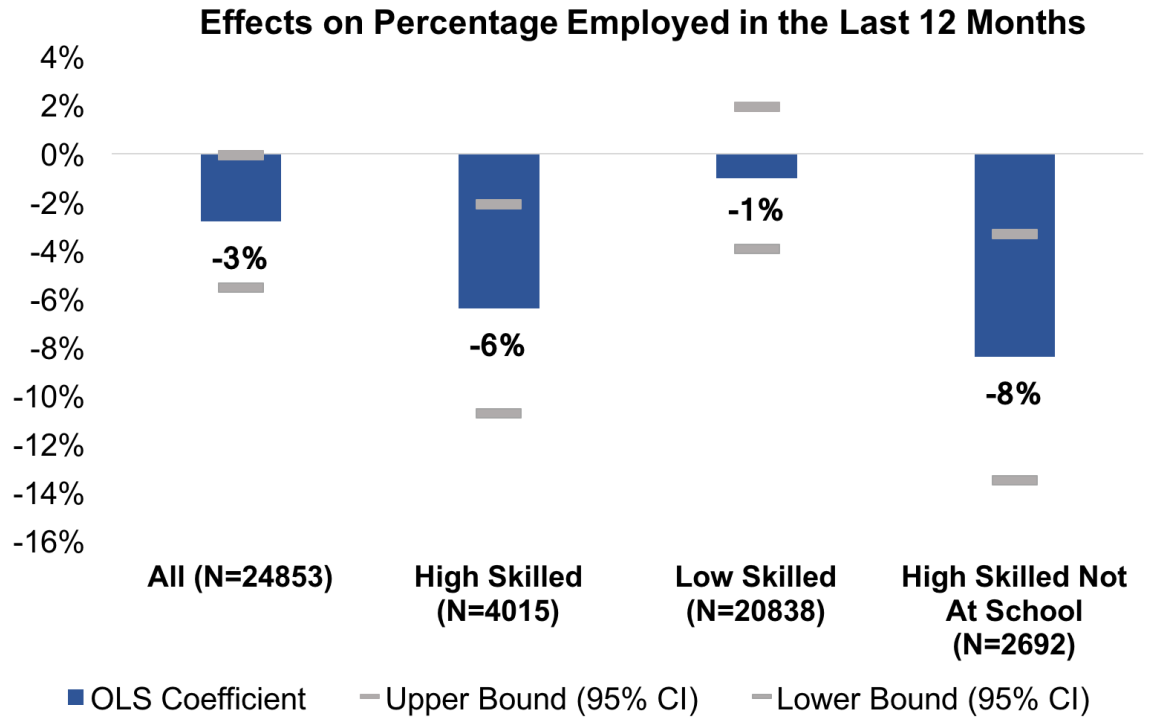
(c) Public Wage Premium at Entry



(d) Post-Secondary Wage Premium

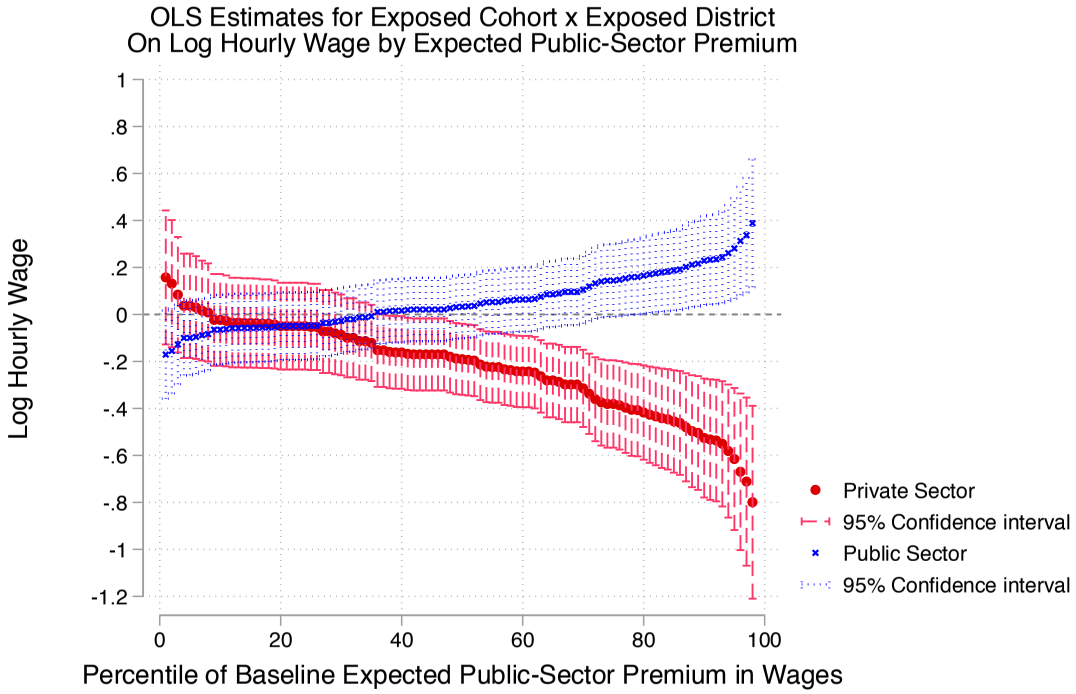
Notes: The figure shows the coefficient on $1[University_d] \times 1[Exposed_{i,c,d}]$ in the regression as per equation 10 for each percentile of the expected public-sector wage premium (top-left). The expected public-sector wage premium is the difference in average public-sector wages and private-sector wages for those aged 25 and older at the time that the nearest public university was established multiplied by the proportion of the post-secondary educated that are employed in the public sector for the same age group. The top-right figure shows the coefficient on $1[University_d] \times 1[Exposed_{i,c,d}]$ for each percentile of the average difference between public wages and private wages for those aged 25 or older at the time that the nearest public university was established. The bottom-left figure shows the coefficient on $1[University_d] \times 1[Exposed_{i,c,d}]$ for each percentile of the average difference between public wages and private wages for those aged 25-28 at the time that the nearest public university was established. The bottom-right figure shows the coefficient on $1[University_d] \times 1[Exposed_{i,c,d}]$ for each percentile of the average difference between log wages for individuals with a post-secondary education and individuals without a post-secondary education for those aged 25 and older at the time that the nearest public university was established.

Figure 9: Heterogenous Effects: Higher Education Attainment by Expected Public-Sector Premium



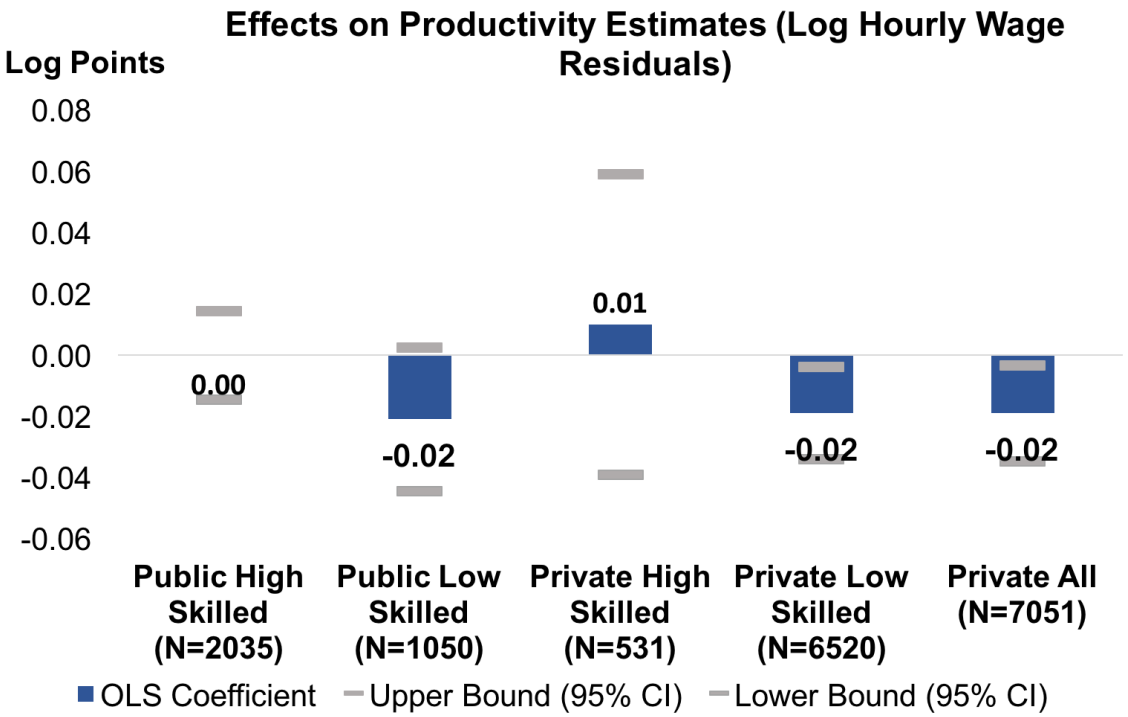
Notes: The figure shows the coefficient on $1[University_d] \times 1[Exposed_{i,c,d}]$ (blue bars) and 95% confidence intervals (grey lines) from a regression of an indicator for whether the individual has been employed over the last 12 months on $1[University_d]$, $1[Exposed_{i,c,d}]$, the interaction, cohort-of-birth fixed effects, district fixed effects and an indicator for whether the individual is a female, as per equation 10. Confidence intervals are calculated using standard errors clustered at the district level. The left bar shows the result for all individuals in the sample of those aged 17-24 at the time the nearest university was established. The second bar shows the result restricting the sample further only to those with a post-secondary education. The third bar restricts the sample to those without a post-secondary education. The fourth bar restricts the sample to those with a post-secondary education and who are not currently at school.

Figure 10: Employment Rates



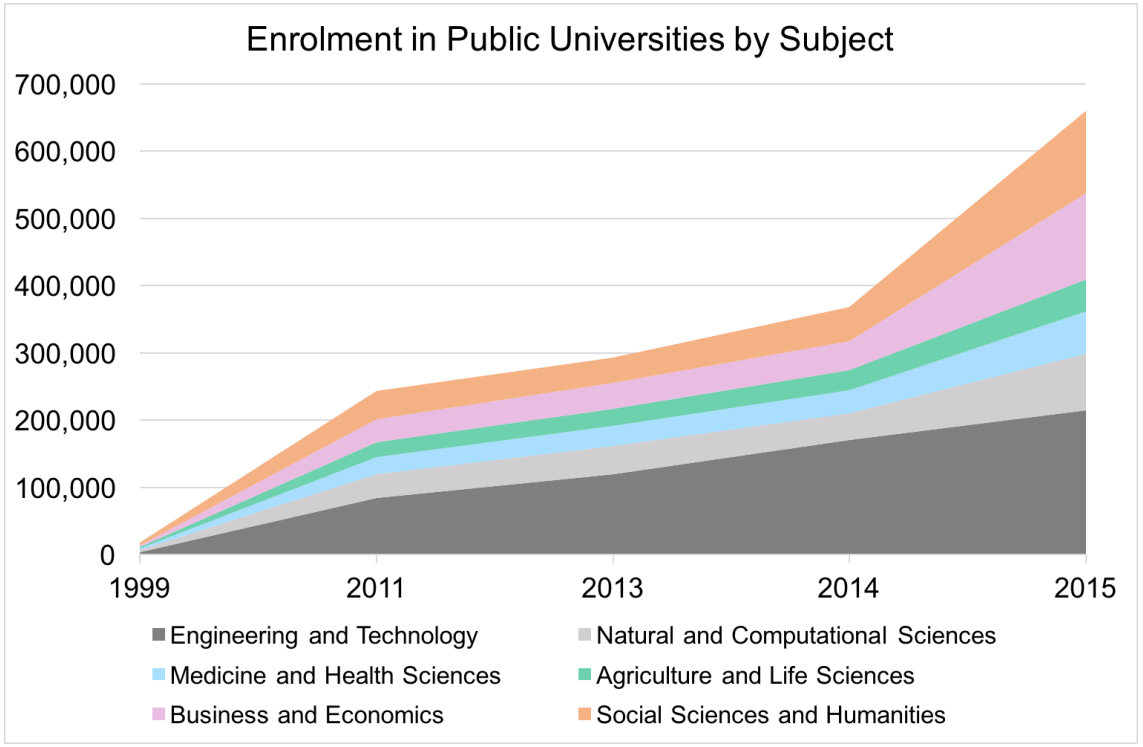
Notes: The figure shows the coefficient on $1[University_d] \times 1[Exposed_{i,c,d}]$ for the private sector (red dots) and 95% confidence intervals (red dashed lines) and the coefficient on $1[University_d] \times 1[Exposed_{i,c,d}]$ plus the coefficient on $1[University_d] \times 1[Exposed_{i,c,d}] \times 1[Public]_{i,c,d}$, the total effect of the expansion for public-sector wages (blue dots) and 95% confidence intervals (blue dashed lines) in the regression as per equation 11 with the dependent variable as the log hourly wage, for each percentile of the expected public-sector wage premium for all unexposed cohorts. The expected public-sector wage premium is calculated as the difference in average public-sector wages and private-sector wages for those aged 25 and older at the time that the nearest public university was established multiplied by the proportion of the post-secondary educated in public employment for the same age group.

Figure 11: Heterogenous Effects: Wage Effects by Expected Public-Sector Wage Premium



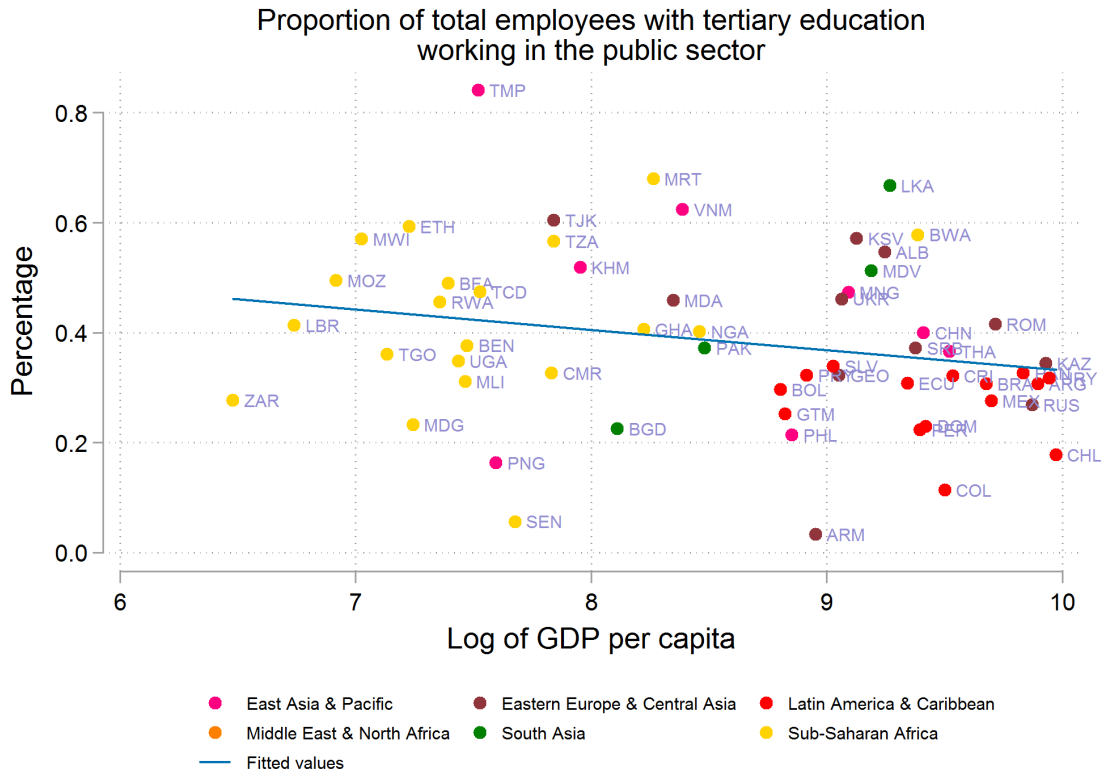
Notes: The figure shows the OLS coefficient (blue bars) and 95% confidence intervals (grey lines) for $1[University_d] \times 1[Exposed_{i,c,d}]$ from a regression of the productivity estimate on $1[University_d]$, $1[Exposed_{i,c,d}]$, and the interaction, for each 'market' separately (public high skilled, public low skilled, private high skilled, and private low skilled), clustering standard errors at the district level. The productivity estimate is the residual from a regression of log hourly wages on an indicator for whether the individual is female, indicators for the broad education level of the individual, cohort fixed effects, and district fixed effects for each 'market'.

Figure 12: Labour Productivity



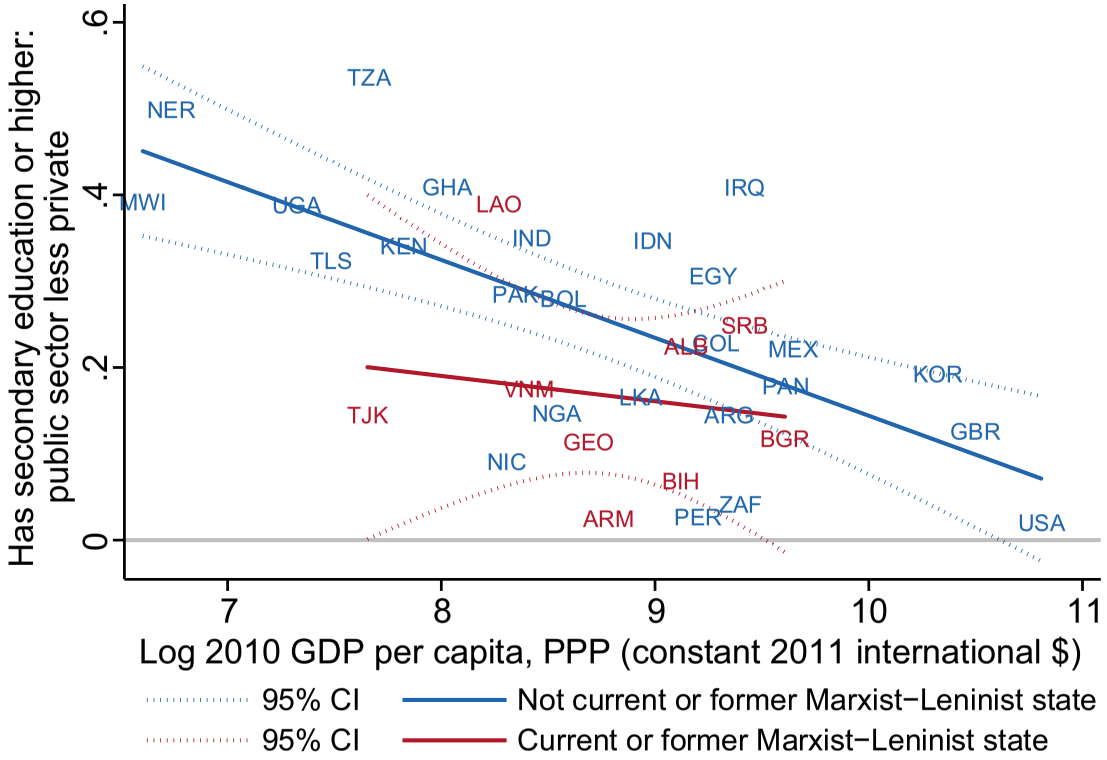
Notes: The figure shows the enrolment in public universities by the six major subject bands of study. The 1999 data is hand-coded to match the six major subject bands. 2011 is the year from which the enrolment by the six major subject bands is publicly reported.

Figure 13: Enrolment in Public Universities by Subject Band



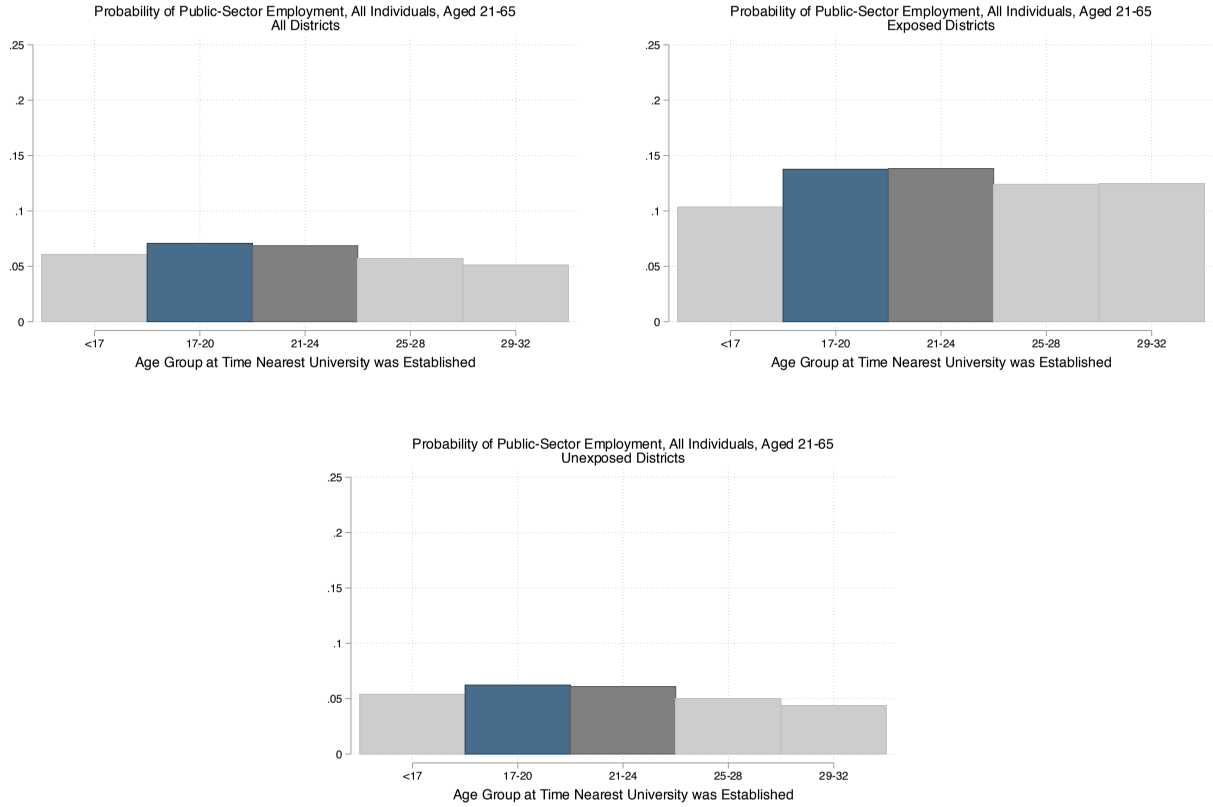
Notes: The figure shows the proportion of employed tertiary-educated individuals working in the public sector. Source: Worldwide Bureaucracy Indicators, World Bank.

Figure 14: Proportion of Employed Tertiary-Educated in Public-Sector Employment



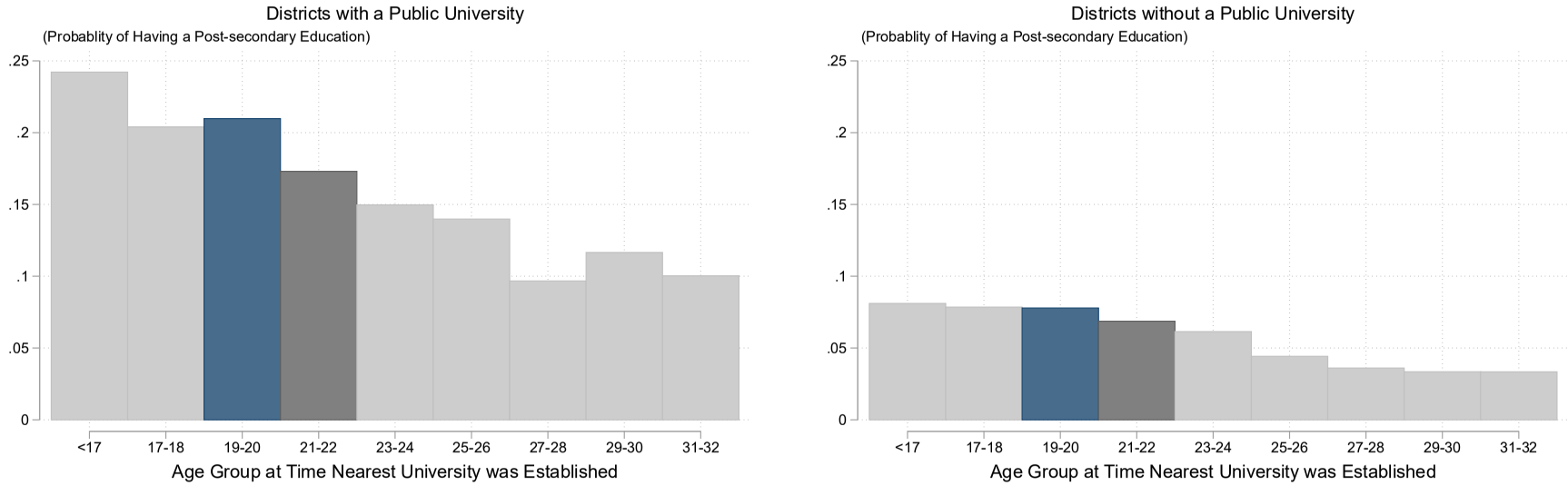
Notes: The figure above is figure 8 from Finan et al. [2017]. The figure shows the proportion of individuals with a secondary education or higher in the public sector minus the proportion in the private sector (the education premium in the public sector).

Figure 15: Secondary and Tertiary Education Premium in the Public Sector [Finan et al., 2017]



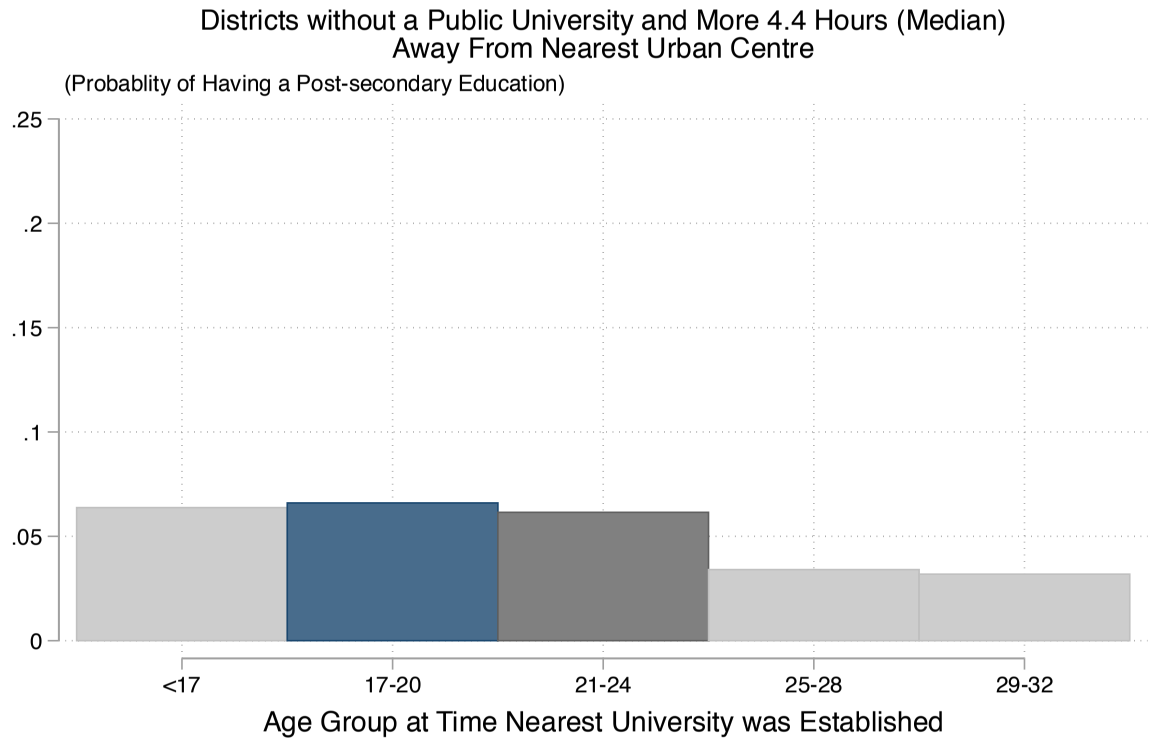
Notes: The top-left figure shows the proportion of public-sector employees from all individuals aged 21-65 by cohort, in terms of age at the time that the nearest public university was established. The top-right figure shows the proportion of public-sector employees out of all individuals aged 21-65 by cohort, for 'exposed' districts (those with a public university). The bottom figure shows the proportion of public-sector employees out of all individuals aged 21-65 by cohort, for 'unexposed' districts (those without a public university). The proportions are calculated using weights based on the inverse of the Labour Force Survey 2013 primary sampling unit population.

Figure 16: Public-Sector Employment by Cohort



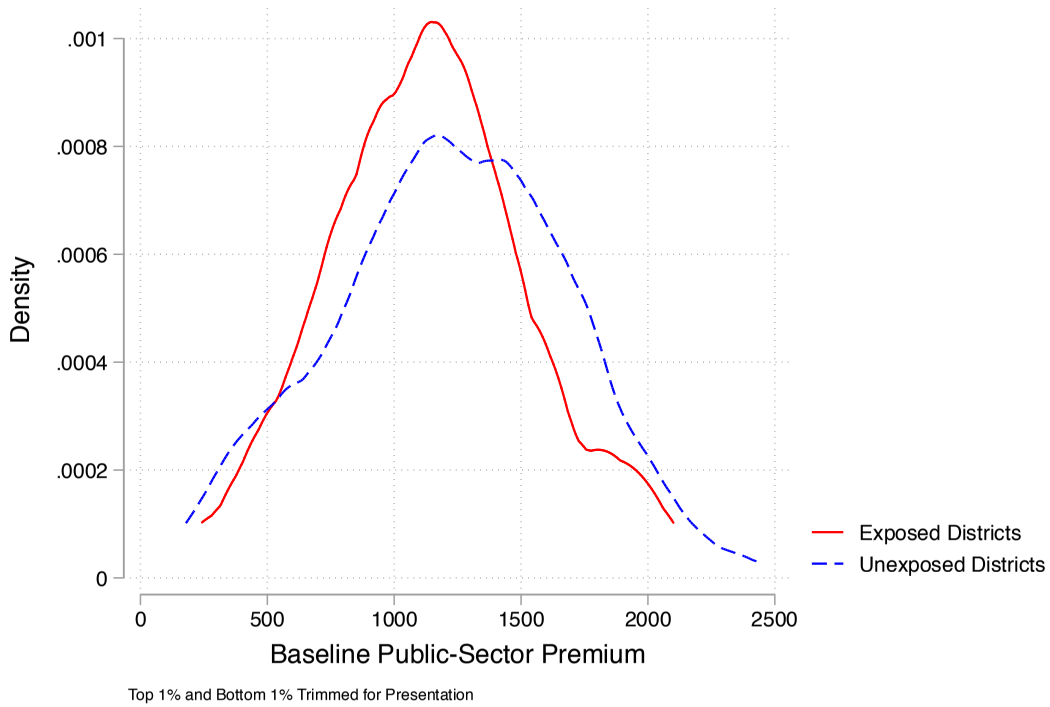
Notes: The figure shows the unconditional probability of having post-secondary education by the age of the respondent at the time that the nearest university was fully established. The probabilities are calculated using weights based on the inverse of the Labour Force Survey 2013 primary sampling unit population. The left figure show those districts that have a public university. The right figure shows districts that do not have a public university.

Figure 17: Education Attainment by Cohort and District Exposure



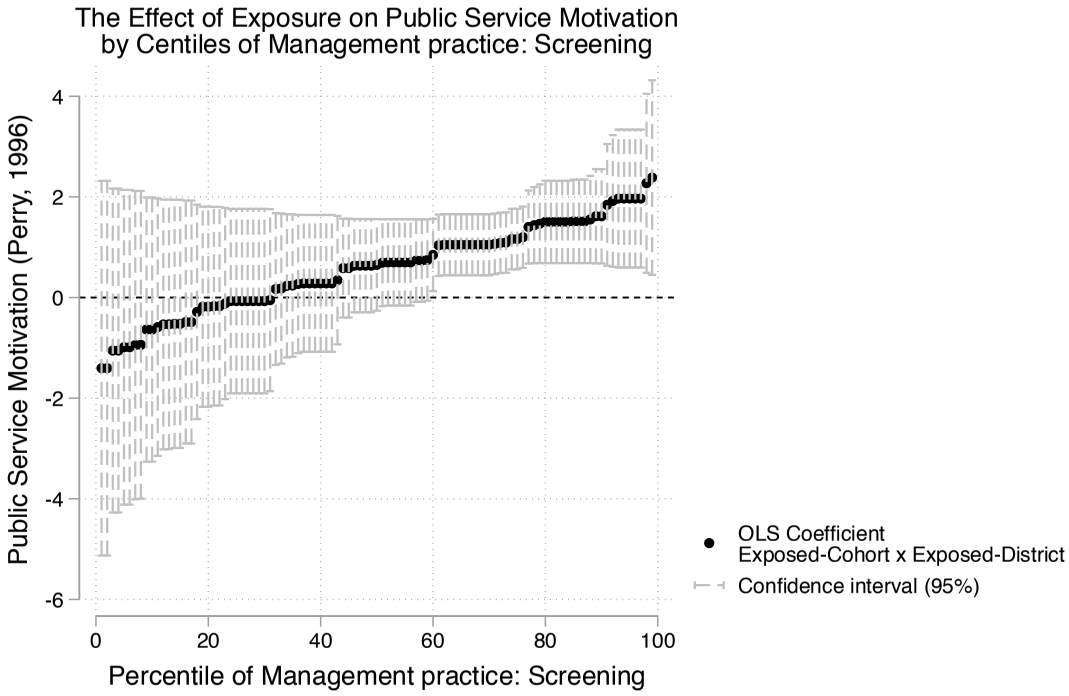
Notes: The figure shows the probability of having a post-secondary education by the age of the respondent at the time that the nearest university was fully established. The sample is restricted to those districts without a public university and at least 4.4 hours away from the nearest public university (the median across districts without a public university). The probabilities are calculated using weights based on the inverse of the Labour Force Survey 2013 primary sampling unit population.

Figure 18: Education Attainment by Cohort: Most Remote Districts



Notes: The figure shows the kernel density, using an Epanechnikov kernel function of the baseline public-sector premium (the average difference between public and private wages in the district for all individuals for the 'exposed' districts (with a public university) in the red solid line, and 'unexposed' districts (without a public university) in the blue dashed line.

Figure 19: Distributions of the Public-Sector Premium in Exposed and Unexposed Districts



Notes: The figure shows the OLS coefficient on $1[University_d] \times 1[Exposed_{i,c,d}]$ and 95% confidence intervals in the regression as per equation 12 with the dependent variable as the z-score of public-service motivation as per Perry (1996), for each percentile of Management practice: Screening

Figure 20: The Effects of University Expansion on Bureaucrat Public-Service Motivation, by Management Practices Related to Screening

3 Hierarchy and Information

3.1 Introduction

The information set of public officials is a key variable of interest in economic policy. The ultimate goal of much research in economics is to inform policymakers, yet we have little evidence on what public officials know and how they absorb new information. In contrast to the classical version of the competitive private sector, where information is acquired at zero cost, the costs and benefits of information acquisition in the public sector are key determinants of the distribution of information. This paper assesses the nature of incentives associated with the acquisition of information within public organizations and determines that these incentives mediate the effectiveness of interventions aimed at improving the information of public-sector agents.

We collect individual-level data on the information held by Ethiopian public officials regarding the characteristics of the local constituents they serve and compare this to objective benchmark data. Our survey covers a representative sample of 1,831 public officials across 382 organizations spanning all three tiers of government. The information variables we collect correspond to fundamental quantities relating to the core work of all of the public officials we interviewed.

Combined with organization-level data on the management practices of public offices and a field experiment, we arrive at four main conclusions regarding the characteristics and determinants of information in bureaucracy. Firstly, officials make large errors about the basic conditions of local jurisdictions. For example, 50% of officials claim that they serve a population that is at least 50% smaller or 50% larger than it is, and across the variables we study, 49% of public officials make errors that are at least 50% of the underlying benchmark data. Secondly, granting formal authority to agents significantly improves the precision of their information by a third of a standard deviation on average. Thirdly, experimentally reducing the cost of access to information also improves the precision of public officials' estimates by around a third of a standard deviation on average. Critically, we find significant heterogeneity in our estimates by management practices. Poorly managed organizations do not respond to the delegation of authority, while the experimental treatment has substantially larger effects for these organizations. This paper, therefore, shows that organizational incentives to acquire information determine the efficacy of interventions and technologies aimed at improving the information of public-sector agents.

These results matter because public officials state that their tacit knowledge is a key source of information in their decision making related to public policy. As Duflo [2017] argues, policymakers "tend to decide on [policy design details] based on hunches, with-

out much regard for evidence." Our survey asks where officials source their knowledge to make public policy decisions. Consistent with other studies we find that bureaucrats rely dominantly on their own, and their colleagues, tacit knowledge. Officials' personal beliefs over the characteristics of the citizens they serve therefore matter for resource allocation decisions and the effectiveness of implementing public policy.

We test this knowledge by asking respondents specific questions on the general socio-economic features of the constituencies they serve, such as the population, and sector-specific features, such as primary enrolment, for the sector in which they work. These variables are fundamental quantities that underlie officials' core work, and as such officials would optimally have knowledge of their values for the districts they serve. We compare these claims to objective benchmark data from administrative data and independent surveys to identify the errors that they make. The errors in their beliefs imply large misallocations of public resources that could be substantially improved by changes in incentives. In a simple scenario in which we model the public officials we study equitably distributing resources across the constituencies they serve, districts at the 75th percentile of errors made by officials receive 30% more funds than their actual population warrants. Compounded by further errors on the nature of citizen need, such as indicators of health status, our data imply that public officials are poorly targeting public resources due to a lack of information.

To our knowledge this is one of the first direct measures of individual-level information in related literatures. Fisher et al. [2018] record 'knowledge-seeking' activities (proxied by downloads from and visits to a particular website) which are distinct to direct assessments of individuals' information on local conditions. There is a body of literature in finance that attempts to measure the ('soft') information held by loan officers at lending institutions and how the acquisition and use of this information responds to incentives. The measures of soft information are either elicited from survey questions that ask directly the extent to which employees feel that they know borrower characteristics (Uchida et al. [2012]; Hattori et al. [2015]), the predictive power of internal risk ratings relative to 'hard' measurable characteristics of firms (Agarwal and Hauswald [2010*a*]; Brown et al. [2012]; Qian et al. [2015]; Agarwal and Ben-David [2018]), or measures of the effort exerted by relationship managers to learn about their clients (Liberti [2017]). We provide direct measures of the knowledge of actors across the hierarchy. Given the novelty of our data and the scarcity of empirical evidence on this subject, we provide a substantial amount of descriptive evidence on the nature and scale of errors in claims made by public officials of Ethiopia's government. If economic research aims to influence the beliefs of public officials, these results provide a baseline assessment of the nature of those beliefs.

In a minority of cases, the public officials we study make relatively accurate claims

about their constituents. Of officials' assessments of the population they serve, 21.5 percent are within 20% of the Census-defined population. However, a large portion of bureaucrats make economically-meaningful mistakes. Roughly half cannot correctly rank the 25th and 75th percentile of districts in terms of population within the region in which they specialize. Such large errors are consistent with studies of agents from other settings, such as farmers not optimizing important aspects of the technology they use (Hanna et al. [2014]), Indian manufacturing firms having limited information on quality defects, machine downtime, and inventory (Bloom et al. [2012]), and professional forecasters making large but predictable mistakes (Ehrbeck and Waldmann [1996]; Andrade and Le Bihan [2013]). In government, such errors can have large economic consequences as they skew the distribution of public resources.

Critical to the study of information in systems of bureaucratic hierarchy is to appreciate the dependence of information acquisition decisions on the corresponding decisions of other agents. If a principal invests heavily in the aggregation and absorption of information, the corresponding agent is disincentivized to undertake costly investments themselves. This is both because they are more likely to be overruled, as well as because they can free-ride on the investments of the principal. Thus, we begin our paper by surveying the related literature on information acquisition in hierarchy, and use the framework of Aghion and Tiróle [1997], henceforth AT, a seminal work in this area, to organize our empirical work around three predictions prominent in the literature.

Our findings are consistent with these predictions. First, the delegation of de facto authority to make production decisions increases the agent's initiative to invest in information acquisition. As AT states, "because the principal cannot overrule the agent, the agent has more incentives to become informed." We use the fact that the Ethiopian government is organized around multiple tiers of government offices to proxy the delegation of de facto authority across tiers, and compare its impacts to the principal-agent relationships of managers and employees within organizations, where principals retain control rights.

Specifically, when testing the implications of de facto authority, we view centralized officials as the principals and decentralized officials as the agents. When testing the implications of de jure authority, we view managers as principals and employees as agents. By limiting our comparisons to claims about the same indicator, at the same percentile of its distribution, for the same district, and controlling for individual factors such as education and experience, we are able to show that providing officials with de facto authority improves the precision of their claims about constituents by a third of a standard deviation. This effect dwarfs the impact of increasing the complexity of the underlying information, such as asking officials to make claims about increasingly poor districts.

The second prediction of the theoretical literature relates to management practices as mediating factors in the provision of de facto authority. In those organizations that have weak incentives for information acquisition, the delegation of authority may have limited impact since there is little impetus for actors to use it. To measure the broad incentive environment of the organizations we study, we follow the methodological approach of the World Management Survey, as laid out in Bloom and Van Reenen [2007*a*] and Bloom et al. [2012]. We adapt their management surveys to the Ethiopian public sector setting, taking account of insights from the public administration literature (Rose-Ackerman [1986], James [1989]). We focus on those management practices that relate to the use of monitoring and thus the value of information within an organization, and those that aim to align the agent to the principal's wishes.

We find significant evidence that management matters for information acquisition by public officials. The delegation of de facto authority has a positive effect on the acquisition of information only when the appropriate management practices are in place, in line with the theoretical prediction. This implies that decentralization's effect on the information absorbed by officials is conditional on appropriate incentives existing at the organization level. District officials working in those organizations with the weakest management practices related to monitoring are no more informed than their central colleagues.

Third, we experimentally test the prediction that a substantial reduction in the cost of acquiring information leads to its absorption. Due to free-riding concerns, in a principal-agent framework, lowering the cost of acquisition of information has an ambiguous impact on the relative distribution of knowledge. We provide an information package that exactly mimics circulars (internal communications) of the Ethiopian government to a random sub-set of the officials that we study. The information package contains all the information we test for in our survey. Lowering the marginal cost of information acquisition dramatically and symmetrically in this way should yield superior information overall. We find an average treatment effect equivalent in magnitude to the impact of improved incentives. However, the treatment effect arises totally from organizations with below-average management practices, reinforcing our findings that organizational practices are key mediating factors in how new information is received and absorbed. This result is of significance to researchers who aim to update the beliefs of public officials.

More broadly, the variation in the precision of public officials' estimates does not seem to be due to differences in the marginal costs of acquiring information beyond the complexity of the underlying variable. Such costs may be mediated by personal characteristics, such as an official's education or experience, or by their proximity to the local environment we are questioning them on. However, the coefficients on demo-

graphic variables and on proxies of the remoteness of the officials from their constituents do not explain a substantial proportion of the variation in errors. The incentives for information acquisition explain substantially more of the variation in errors we observe.

In line with much of the recent literature on bureaucrats (such as Dal Bó et al. [2013] and Bertrand et al. [2016]), our focus is to understand the determinants of the characteristics of bureaucracy, rather than its wider effects. However, we find some evidence that superior information improves promotion prospects of bureaucrats and, in the form of stating knowledge of the indicators we focus on, improves service delivery outcomes. This is consistent with descriptive evidence from our survey that information held by civil servants on the conditions in the constituency they serve feeds into allocation and operational decisions. Officials in our survey state that on 74% of projects, "information regarding the state of operations and services in the jurisdiction makes it easier for me to know where to allocate my effort/time". Similarly, 82% of officials stated that jurisdictional information matters for a majority of the operations (projects/programs) that they work on. It is also consistent with Callen et al. [2018] who document increases in doctor attendance from a system that highlights poorly performing districts in Pakistan to Executive District Officers. At the same time, the scale of the errors does not systematically predict the service delivery indicators we have information on. Neither does detailed information on the absence rate of doctors provided to local officials in India [Dhaliwal and Hanna, 2017]. Together, our results highlight the need for more research on how the information of public officials and its equilibrium distribution in government relate to the delivery of public goods.

Our experimental findings highlight that when the public sector is designed appropriately, the rationale for compensating interventions - which are often ad hoc and focused on short term goals - is weakened. The results echo those of Olken [2007] in which top-down government monitoring reduces corruption substantially more than grass-roots monitoring, Muralidharan and Sundararaman [2011] where effectively designed incentive schemes improved student learning outcomes far more than unconditional grants, and Pomeranz [2015] in which pre-existing incentives for tax compliance substitute for government audits. While external interventions can reduce the scale of errors made by public officials, suitable organizational incentives ensure self-acquisition of that information. The paper thus relates to the nascent literature on how the design of bureaucratic organizations impacts on state capabilities (reviewed by Finan et al. [2017]), and on the importance of these capabilities for development outcomes (Besley and Persson [2011]; Acemoğlu and Robinson [2012*b*]; Pepinsky et al. [2017]).

This paper also contributes to the literature on public sector decentralization. We provide direct empirical evidence that decentralized organizations have superior information, which is often assumed in the literature [Oates, 1999; OATES, 2005; Bardhan

and Mookherjee, 2006]. Our evidence implies that within layers of an organization, information is relatively homogeneous, while across organizational tiers it is relatively asymmetric. Our finding that the impact of decentralization on information is mediated by management practices provides one potential explanation for the divergent findings of the decentralization literature on service delivery [Bardhan, 2002; Crawford and Hartmann, 2008]. The rest of this paper proceeds as follows. The next section describes the related literature on this topic and the predictions derived from the theory. The 'Measuring Information' section details the approach we took to measuring information in the public sector. The section 'The Determinants of Superior Information in Hierarchy' presents detailed descriptives of the determinants of information in hierarchy. 'Experimental Evidence on Information in Hierarchy' outlines our experimental design and presents results. The final section concludes.

3.2 Literature Review

In much of contract literature, information is assumed to be an exogenous endowment of agents.[22] Argenziano et al. [2016] and Di Pei [2015] argue that studying endogenous information acquisition is key to understanding the distribution of information among groups of players. Both papers show that when information gathering is costly, the interaction between moral hazard in acquiring information and adverse selection in reporting information overturns core predictions in Crawford and Sobel [1982] paper, such as there being a greater degree of information sharing when preferences are closely aligned. In a review of almost 50 years of research on information economics, Stiglitz [2017] underlines the importance of recognizing the endogeneity of information imperfections.

Aghion and Tiróle [1997], henceforth AT, was an early attempt to endogenize the information acquisition process, and it acts as a useful framework for organizing the literature on the topic. AT highlights that the acquisition of information is costly and thus there must be an incentive for its collection and absorption. In environments of collective output, such as hierarchical bureaucracies, this implies that officials may free-ride on each other's efforts for information acquisition.

AT assume a principal and agent jointly determine the features of a project about which the nature of payoffs is initially unknown and characteristics of which cannot be contracted on. With costly effort e_i, an actor is fully informed about these payoffs with probability $e_i \in (0, 1)$, but with probability $1 - e_i$ their investment yields no new information. The incomplete contract between the two parties specifies an allocation

[22]Though we define our empirical conception of information precisely in the following section, while discussing the related literature we follow Mookherjee [2006] definition of information as 'the knowledge of agents relevant to the efficient allocation of resources'.

of formal authority, which they define as the right to choose project particulars, to one of the two actors.

When an individual i has control rights, payoffs are of the form:

$$u_i(e_i, e_j) = e_i U_{ii} + (1 - e_i)\, e_j U_{ij} - g\,(e_i) \tag{16}$$

where the utility of player i is a function of the probability i is informed, e_i, the utility from a successful project if i has control rights, U_{ii}, or if j has control rights, U_{ij}, the probability that the other actor is informed, e_j, and the cost of investment. Otherwise, payoffs are conditional on the decisions of the other player.[23]

One can immediately see that there is a tradeoff in the second term of (16) between investments by the principal and those by the agent. By lowering investment in information acquisition when e_j is high, they can free-ride on the efforts of their colleague to identify the payoffs of project choices. This simple setup leads us to three core predictions that clearly summarize insights from across the field.

(i) Prediction 1: The transfer of formal authority to an agent increases the agent's incentive to acquire information.

In the case in which the principal has control rights, the agent provides a signal of their information if informed to the principal, and the agent's utility is determined by how congruent the principal's preferred project is to their preferences. Under principal-authority, when the agent does not have control rights, the first order condition of (16) for the agent is,

$$g'\,(e_i) = U_{ii} - e_j U_{ii}$$

while under agent-authority, the agent makes the decision over project particulars and the corresponding condition is,

$$g'\,(e_i) = U_{ii} - e_j U_{ij}$$

Since in general U_{ij}, the benefit to the agent if the principal has control rights, will be smaller than U_{ii}, the agent invests greater effort in information acquisition when they have control rights. These first order conditions are equivalent to the production function for knowledge in Garicano [2000]. Dessein [2002] finds the same result in a model in which the nature of authority is endogenous.

This provides a theoretical rationale for the assumption that decentralized agents

[23] If both players are uninformed, AT assume no actions are taken. The cost function is increasing and convex. An appendix presents a fuller description of the AT model and our corresponding derivations.

have superior information on local conditions (recent examples of which Acemoglu et al. [2007] and Aghion et al. [2017]). It highlights that the nature of control rights, rather than hierarchical distance from the principal or proximity to the lowest level of hierarchy, is a key determining factor in the acquisition of information.

Closely related to these insights are those of the political economy literature, in which decentralized agents (such as local governments in a multi-tier system) are assumed to have superior information due to their proximity to local conditions. As Oates [1999] states, "individual local governments are presumably much closer to the people and geography of their respective jurisdictions; they possess knowledge of both local preferences and cost conditions that a central agency is unlikely to have". Bardhan [2002] argues that the contracting environment of local governments incentivizes information acquisition through political accountability or yardstick competition, and that the cost of access to information is lower relative to central level governments, implying differential $g(\cdot)$ functions. AT and the related literature argue that superior information at decentralized levels may be an equilibrium outcome dependent not on proximity to the locale, but on control rights. Our empirical results provide little support to the theories of proximity and substantial evidence of the importance of control rights.

Cremer et al. [1994] discuss the implications of applying this thinking to public service delivery, and emphasize the importance of an incomplete contracts framework to rationalizing any information asymmetries between tiers of government. As they argue, in acquiring information, "there is no reason why central government cannot use any of the techniques available to local government", such that it must be an incentive issue.[24] Similarly, contract theory aligned with Myerson [1981] and summarized in Mookherjee [2006] has pointed out that if information is 'hard' then under the Revelation Principle there is no reason to expect formal delegation to lead to superior information. Since under credible commitment and without constraints on contract complexity and information exchange, the central organization can perfectly mimic the delegated organization. As Petersen [2004] argues, the crucial attribute of models of information in hierarchy is that information assets are subjective in nature, such that contracts must be incomplete. Much of the recent contract literature has focused on the differential implications of hard versus soft information, especially in financial intermediation and credit decisions Stein [2002]. As we will see in section 3.3.3, much of the knowledge in the public sector is of a tacit nature and thus uncontractable, such that the 'soft' class

[24]Cremer et al. continue, "Central governments do name representatives to local areas who collect information on their behalf (for instance, the French prefects); and they do use mechanisms by which citizens of a local area can express their views and provide information. If anything, the balance of advantage as far as the availability of information gathering techniques is concerned probably lies with central government. There are economies of scope in the collection of information: a central agency which runs many transportation systems can transfer techniques acquired in one of them to the others at low cost; it can also afford to hire technicians in more specialized areas than can a local agency."

of models is the most appropriate for our setting, and our framework should be one of incomplete contracting.[25]

There have been no direct tests of these theories of information across hierarchy that we know of, nor has there been much empirical investigation into metrics of what constitutes the boundary between hard and soft information in a particular setting. An attempt at the latter is Agarwal and Hauswald [2010*b*], who define hard and soft elements of credit decisions, and study which are taken by decentralized branches of a bank. Similarly, Liberti [2017] finds that delegation of formal authority empowers loan officers to acquire and use soft information in lending decisions.[26] Overall, whether formal delegation generates superior information regarding local operating conditions is an open empirical question.

(ii) Prediction 2: Increasing the benefit of being informed unambiguously increases information of both the principal and agent irrespective of who has formal authority.

Returning to equation (16), the utility from acquiring information, U_{ik} $(k = i, j)$, increases the incentive to invest in information acquisition under all forms of authority. Importantly, the scale of U_i will play an important role in the responsiveness of actors to changes in control rights. These results are in line with predictions from Gant et al. [2002] that human resource practices, a proxy for the incentives under which actors work, are central determinants of whether agents undertake decentralized problem solving.

Mookherjee [2006] makes the point that the value of incentive regimes such as management practices that foster individual initiative in information absorption "cannot be explained by any conventional incentive-based theory of adverse selection or moral hazard that conforms to the Revelation Principle". If the Revelation Principle could be applied, any decentralized mechanism could be replicated by a centralized mechanism in which central officials manage each agent directly to aggregate all available information. Thus, the impact of management practices on information acquisition must be working through their mediation of the corresponding incomplete contract.

Qian et al. [2015] investigate a reform in Chinese banks that delegated authority to individual loan officers, as a plausibly exogenous shock to loan officer incentives to

[25]The closest to 'hard' information in our setting is bureaucrat knowledge of the Ethiopian government's public service rules. We test for whether claims related to the public service rules follow the implications of a 'hard-information' model in which information can easily be centralized in section 3.4.2.

[26]The measure of soft information in the analysis is the effort of the relationship manager (RM) within the bank in terms of time spent on accounts, the attitude of the RM, and the general satisfaction with the RM.

produce information. They find that the soft information of the bank, as measured by the extent to which their internal risk rating predicts loan interest rates, improves after the delegation. The authors conclude that incentives affect information production and use. Recent empirical evidence has shown that management affects the distribution of assets across firms such as human capital [Bender et al., 2018; Bloom and Van Reenen, 2007*a*] and physical capital [Bloom et al., 2010]. It therefore seems plausible that it would have substantive impacts on the distribution of information.

An additional insight from AT is related to the role of alignment of preferences between principals and agents. Differentiating the utility functions in AT corresponding to (16) with respect to their alignment parameters, we see that an increase in alignment around the equilibrium levels of effort reduces information acquisition by the actor who has control rights and increases it for whoever does not have control rights. With control rights, an official can free-ride on the choices of aligned colleagues, while those colleagues expect this and increase their investment in information acquisition. This is in contrast to the seminal work of Crawford and Sobel [1982], who show that the transmission of information improves as the preferences of the agent and the preferences of the principal become more aligned. However, it is in line with Myatt and Wallace [2011], who show that the number of signals sent shrinks as the complementarity between actions increases in a coordination game with endogenous information acquisition. Hence the association between preference alignment within the organization and the distribution of information in equilibrium is theoretically ambiguous, and depends on the institutional environment.

(iii) Prediction 3: A symmetric decrease in the marginal cost of information acquisition increases the information of both the principal and the agent.

Totally differentiating the first order conditions of the core equations of AT with respect to g' (see theory appendix for details), we see that a simultaneous and *symmetric* reduction in the marginal cost of investment for the principal and the agent will lead to an increase in investment in the acquisition of information.[27] This finding is echoed in Gabaix et al. [2006], where information acquisition is a decreasing function of both the cost of acquisition and its complexity, and Garicano [2000], in which cheap information is widely shared within an organization. Myatt and Wallace [2011] study how much costly attention players in a coordination game should pay to signals that vary in

[27]Reducing the marginal cost for the principal (agent) only will increase the investment of the principal (agent) but decrease the investment of the agent (principal) due to the strategic substitutability between the two. The total differentiation leads to this prediction under the equilibrium stability conditions imposed in AT. By 'symmetric' we refer to the total investment response of agents as a result of the reduction in costs, since the responses will depend on the equilibrium values of effort and the shapes of the cost functions. See the Appendix for further details.

accuracy (how precisely it identifies the state) and clarity (how easy it is to understand), finding that the clarity of signals is of primary importance even if they have poor underlying accuracy.

Empirically, there is evidence from a range of settings that reducing the cost of becoming informed increases information acquisition, as one might expect. Hastings and Weinstein [2008] find that providing a simple information packet to parents regarding local school test results increases the proportion of parents that factor this information into their school choice. Jessoe and Rapson [2014] find that electricity consumers become more responsive to price changes when presented with high-frequency information about their residential electricity usage through a simple in-home display. Banerjee et al. [2018] show that mailing Indonesian households information on their eligibility for rice subsidies resulted in a significant increase in the amount of subsidy that they received, implying that they absorbed and acted on the information. In the Banerjee et al study, the more information that was provided on the mailing, the larger the treatment effect, implying that it was the information, rather than simply the presence of a reminder that was driving the results.

It is unclear, however, that these results on decision making by atomistic households translate to the public sector setting. Given the wider free-riding dynamics of information in hierarchy, it is an empirical question as to whether even large decreases in the cost of acquiring information have impacts on its equilibrium level. We therefore extend this empirical literature by studying the incentive environment around information acquisition to understand its determinants.

3.3 Measuring Information

3.3.1 Institutional Background

Ethiopia is Africa's second most populous country, home to 100 million people or 10% of the population of Sub-Saharan Africa. Its ethnically diverse population is dispersed across over a million square kilometers, such that population characteristics vary widely across space.

Like other developing countries, government expenditures represent a significant fraction of GDP (18%) and the public sector is a large employer of formal workers. Corruption is less prevalent than in most of Sub-Saharan Africa, but government effectiveness is seen as relatively weak overall. Never fully colonized, Ethiopia's government is organized as a developmental state, with a relatively vertical and rigid hierarchy.

The state is organized in a federal system, with three major tiers of government: federal, regional, and district (woreda). Each tier has its own staff of civil servants who must define and refine government policy within their tier, direct budgetary and

other resources, and manage the implementation of public services. Though the specific mix and focus of these tasks differs across tiers, each tier undertakes these tasks to a degree, and must make choices informed by the information they have on citizen needs. For example, roughly half of all public resources are spent by the Federal Government, with the rest being directed by the regional governments or spent directly by district administrations [World Bank, 2016*a*]. None of the officials we study implements public services themselves, at any of the tiers, but rather administers public policy and resources.

Across sectors, the structure of the vertical hierarchy is relatively constant. Ministry organizations focus more on agenda setting for the sector, and the development of governing policies, but also direct resources to service providers and thus focus on managing implementation. Regional organizations develop policies for their region in line with federal stipulations, and guide resources across districts. District governments focus on service delivery, while translating policies into guidelines for their district. Each tier therefore makes decisions on how to distribute resources depending on where they believe there is most significant need. This is true in both the development of policy, which dictates the areas of the sector that should receive resources, as well as in implementation, which assigns resources geographically. Federal and regional organizations co-ordinate closely in their development of policy and operational activities, and can be seen as a coherent 'centralized' body that provides joint guidance to district governments.

We use design features of the Ethiopian government that correspond to key elements of the theoretical literature to investigate the distribution of information across the hierarchy. The most fundamental design features of a federal bureaucracy, like that in Ethiopia, are the vertical relationship between tiers of government, and the vertical relationship between managers of organizations at those tiers and their employees. Both exhibit classic features of a principal-agent relationship: co-production, de jure power of the principal, and potentially asymmetric preferences.

The relationship between central government agencies and district governments is such that many basic decision rights lie with the district government. Day-to-day decisions made by members of a district organization cannot be overruled by the central government. The relationship between central and district government agencies therefore accords closely to AT's notion of de facto 'agent authority'. In contrast, within an organization (an 'integrated' environment in the words of AT) a manager has the ability to overrule the agents she employs on business decisions. The relationship between managers and employees therefore accords closely to AT's notion of de facto 'principal authority'. By contrasting the differences in information between principals (central tiers of government or managers) and agents (district governments or employ-

ees) under these two regimes allows us to assess the impact of the nature of authority on information acquisition.

We can also use the fact that different organizations in the Ethiopian context are managed differently (as will be shown below) to investigate the impact of environments that differ in terms of how they value the acquisition of information. As the culture of an organization becomes more empirical, we can interpret this as an increase in the individual rewards for information use, or the U_i's outlined above. We can thus test the impact of mediating incentives on the likelihood of information acquisition.

3.3.2 Survey of Public Officials

The implementation of any of the three core tasks of public administration outlined above requires information on the characteristics and thus needs of the citizens they serve. To elicit an understanding of the information public officials have about their environment and to understand the incentive environments in which they operate, between June and September 2016 we held face-to-face interviews with 1,831 public administrators in 382 organizations across the three tiers of Ethiopia's government. Table A1 lists the organizations included in our study, and Figure A1 provides a map of the district governments we sampled. Our analysis focuses on the agriculture, education, health, revenue, and trade sectors only. For each of these five sectors, we sample the main ministry, all the corresponding regional offices, and a geographically representative sample of corresponding district offices from across the country.[28] Within each organization, we interviewed senior managers and a representative sample of their staff. We limit our scope to the professional grades of technical and administrative officers, excluding grades that cover cleaners, drivers, secretaries, etc.

Table 1 provides descriptives for the 382 organizations we study. All officials work within a relatively standard structure, with a manager overseeing levels of hierarchy below him within a clearly defined organizational structure. As is common, the more centralized an authority is, the larger it is in terms of both the number of managers and the number of employees. Federal organizations also have a higher number of employees per director than either regional or district government offices.

In terms of bureaucrat characteristics, around 20% are women (at managerial grades, 8% are women, at non-managerial grades 24% are women); 82% of Ethiopian bureaucrats have some form of university education, with 11% having a postgraduate degree (at managerial (non-managerial) grades, 17% (10%) have a post-graduate degree). As in other state organizations, bureaucrats enjoy stable employment once in service: the

[28] Our survey, and therefore analysis, necessarily excluded the majority of the Somali region for security reasons.

average bureaucrat has 13 years in service, with their average tenure in the current organization being 7 years.

Across tiers, bureaucrats are similar in that they are in their mid-thirties, have been in their current post for roughly similar amounts of time, and have all worked in roughly 3 service organizations. However, officials at regional organizations are more senior and more likely to be a manager. Officials at district organizations are less likely to have a university education, and only 1% of district officials have a master's degree. They are also more likely to be male, with 18% of district officials female compared with 23% of regional officials and 27% of federal employees.

To develop and enumerate a questionnaire that was relevant for the Ethiopian setting, we worked closely with the Ministry of Public Service and Human Resource Development and employed ex-civil servants within our enumeration teams to facilitate navigation of the public service. As will be described in more detail below, we used the survey to ask a range of questions about citizen characteristics and the public service rules to capture the knowledge of public officials with regards their working environment. Beyond our focus on collecting bureaucrat knowledge, we surveyed officials on their basic characteristics and the nature of management practices utilized at their organization. The implementation of the survey was successful across the organizations we visited, with 99.5% of public officials sampled agreeing to be interviewed and 98.2% of interviews being classified by the enumerator as having gone 'somewhat well' (26.4%) or 'very well' (71.7%).

In conjunction with the survey, we also collected data on the performance of public officials as assessed in their annual appraisal. Each year, public officials are evaluated by their direct manager on the tasks that they were expected to contribute to. For example, tasks might include 'Monitor and provide support to the [work] team preparing the budget' and 'Support the team to prepare soft and hard copy documents of the budget'. Managers evaluate the quality of contributions bureaucrats make to the tasks they were involved in and produce an overall 'performance' score. In addition to this performance-related score, public officials are evaluated on their 'attitude' to work, which intends to measure their office behavior and alignment to the organization [Abagisa, 2014; Tereda, 2014]. For the year 2016, we collected the performance, attitude, and total scores (which are a weighted average of performance and attitude scores) for each official from a subset of the organizations we visited for which they were available.

3.3.3 What is the Appropriate Conception of Information in the Public Sector?

To effectively measure information in the public sector, we must identify the appropriate conceptualization of the concept for this setting. AT emphasize a conception of information in which agents become informed individually through costly investment in knowledge accumulation. [Mookherjee, 2006] defines information as 'the knowledge of agents relevant to the efficient allocation of resources'. This fits our empirical findings from the Ethiopian public sector well.

In our representative survey of public officials, we asked public servants, "What are the three most important sources of information you use to find out about the state of service delivery in your jurisdiction?" along the following margins: 'Informal interactions e.g. through co-workers', 'Management Information Systems', 'Formal field visits', 'Reports from the frontline', 'Information from higher-tier institutions', 'Formal citizen/user interactions', or 'Media sources'. Officials were asked to rank their top three sources or specify that others are more important.

In our setting, information is dominantly sourced from inside the public service rather than externally from the media. Only 8.93% of public officials in our survey responded that media sources were one of their top three sources of information, and 2.12% of officials selected it as their top source of information. This finding was consistent across tiers, with 8.54% of federal, 8.56% of regional and 9.18% of district officials stating media was one of their top-3 information sources, across sectors and across grades. That public officials are heavily reliant on information produced by the public service is consistent with the findings of AidData's 'Decoding Data Use' project [Masaki et al., 2017]. Comparing the sources of information of government officials with actors in the private sector, civil society, independent experts and others, AidData find bureaucrats the least likely to consult external governance assessments in their work. Similarly, a review of Ethiopia's 'Sector Development Program IV: 2010-2015' documents for the five sectors studied in this paper finds almost no references to external sources.

The most frequently cited source of information was 'Formal field visits', with 63% of officials stating that this was a key source of information. Discussions with frontline colleagues, and informal interactions with colleagues were the second and third most cited source of information, with 51.9% and 45.9% of officials stating their significance respectively. A similar pattern is observed restricting the selections to the top cited source of information for each individual: formal field visits are most frequently cited (22.3%), followed by informal interactions with colleagues (19.6%) and then reports from the frontline (17.2%).[29]

[29]Only 4% of officials stated that one of their top three sources of information was not included in

These responses indicate that the dominant form of information used for decision making in the Ethiopian public sector is in terms of knowledge transmitted or gained through personal experience, widely known as the 'tacit' knowledge of individuals. Together, information gained through the direct experience of the individual - from field visits or interacting with colleagues or citizens - is one of the top three sources of information collection for 90% of the officials in our data. Only 12.8% of officials state that they use management information systems (MIS) as their primary source of information. Field visits and informal interactions are 3 times more likely to be in the top 3 most important sources of information than MIS, and 10 times more likely than external media sources.

The reliance on tacit knowledge is in contrast to the fact that many public officials believe relevant information is available from other sources. We asked officials how long it would take to receive statistical data as an input to their work from the relevant directorate, and the median response was 1 day (the mean was 1.8 days and the 90th percentile was 2 days).[30] These findings for Ethiopia's public service are consistent with wider findings from the World Development Report 2016 that highlighted how public officials across the developing world fail to capitalize on digital databases available to them.[31]

This lack of use of MIS does not seem to be due to a lack of skills. Our survey asked the extent to which officials agreed with the statement, "I possess the necessary skills to make best use of the technology that is available to me." Two-thirds of respondents stated that they 'slightly' or 'strongly' agreed with the statement. Similarly, education indicators explain less than 1% of the variation in the probability that officials state that they use informal interactions in their top 3 most important sources of information. The lack of use of MIS also does not seem to be due to variation in facilities such as hours of electricity an organization typically receives, the quality of the phone network or internet access. An aggregate index of the quality of facilities at the organizations we study also explains less than 1% of the variation in the likelihood of relying on tacit knowledge.[32]

---

our survey options.

[30]Averaging bureaucrat responses to the organization level, the 90th percentile slowest organization delivered the data within 3 days.

[31]Many public officials believe that the tacit information they have is sufficient for the successful implementation of their daily work. Our respondents stated that in 70.9% of the projects they worked on, "I have necessary information regarding the current state of service delivery in my jurisdiction/my area of work". This was consistent across tiers (federal, 73.6%; regional, 72.9%; district, 69.4%), sectors (Agriculture, 71.1%; Education, 75.3%; Health, 74.3%; Revenue, 66.1%; Trade, 67.6%) and grades (Grades 1-5, 71.1%; Grades 6+, 70.7%).

[32]The index of facilities is an organization-level average of z-scores of the facilities questions asked to managers only: how many hours of electricity an organization receives per typical working day, the state of the phone network, the extent of internet access, the access staff have to computers, the access

3.3.4 Subjective Assessments of Public Officials

Our survey evidence points to the appropriate conception of information for this setting to be the public official's tacit knowledge of their local environment, or what officials and their colleagues directly know. Such data are not systematically collected by civil service authorities. We therefore used the public official's survey to ask each official direct questions on the socio-economic conditions of the citizens they serve and on civil service rules that govern their workplace. Respondents had to answer based on their existing knowledge of their constituencies and workplace rather than by referring to external reference documents. This process is closest to the conception of knowledge outlined by public officials as being applied in their daily work, as summarized above.

The questions we asked are listed in Table A2 and relate to both district characteristics and civil service rules. Those relating to socio-economic characteristics of citizens correspond to the core variables of interest in national sector plans such as the 'Education Sector Development Program IV: 2010 - 2015'. Those relating to the public service correspond to the rules outlined in the 'Ethiopian Federal Civil Servants Proclamation No. 515/2007'. Thus, all the variables we asked about are fundamental quantities relating to the core work of the public officials we interviewed. For example, the core mission of education officers is to implement the education sector development programs, focusing on the topics it prioritizes.

There are three sets of questions. The first set asks basic demographic questions about districts such as population and the proportion of the population that lives in rural areas. These questions are asked to respondents across all sectors. The second part asks sector-specific questions enumerated to sector-officials only. For example, respondents working in the education sector were asked the following questions about the citizens in their constituency: "What do you think the primary (1-8) enrolment numbers are for [district name] according to official administrative data?"; "What do you think the primary (1-8) pupil-teacher ratio is for [district name] according to official administrative data?"; "What do you think the primary (1-8) pupil-class ratio is for [district name] according to official administrative data?"; and, "What do you think the primary (1-8) pupils-per-school ratio is for [district] according to official administrative data?"[33] The third set ask questions about the rules of employment of the civil service, such as the number of hours a public official is expected to work in a regular working week, the annual leave allowance in days in the first year of service, and so on. These are once again asked to all officials. All indicators asked about were continuous in nature.

staff have to vehicles, and the staff's ability to write memos/spreadsheets/powerpoint.

[33]District-level sector-specific information was not available for the Addis Ababa City Administration, Dire Dawa City Administration and Harar Regional Administration.

Since the three tiers of government serve constituencies of differing sizes, we asked federal officials which region they knew best, and then asked them questions about the 25th, median and 75th percentile districts within that region by socio-economic indicator. Regional officials were asked about the corresponding 25th, median and 75th percentile districts within the region where they worked. District officials were asked about the local government in which they worked. Thus, all officials made statements about specific districts in their constituency, and in particular for the federal officials, those they stated they knew best.[34]

For each question, public officials were permitted to state, "Don't Know", though the enumerator did not suggest this as a response. In 23% of questions, the respondent stated that they simply did not know and could not provide a valid guess. This is in contrast to the fact that the quantities assessed are the core variables required to make policy, planning and distribution decisions in accordance with the national sector plans.

3.3.5 Objective Assessments of Citizen Characteristics

For each indicator, we have administrative or survey data that can benchmark the responses by officials. The specific sources for each variable are outlined in Table A2 and are the latest Census, the latest available service delivery data from the education and health management information systems, and the Woreda and City Benchmarking Survey (WCBS). Each data set is created by sourcing information directly from service providers and aggregating it at the center of government. The sources we use are generally regarded as the highest quality large-scale measurement undertaken in Ethiopia on the topics outlined above. For example, in education, the Education Management Information System is comprised of school-level data that are processed through each tier of government, officials of which undertake quality checks to ensure that the data are as accurate as possible. The WCBS data was collected in collaboration with the World Bank and the Ministry of Public Service and Human Resource Development as a monitoring tool over five separate waves of data collection. All of this data is available to the officials we study through the relevant directorates of their tier of government at the minimum, and in some cases online.

One proxy of the quality of the benchmarking data is their stability over time. By assessing the extent to which indicators vary within districts over data collection rounds, we aim to understand whether the data reflect a noisy collection process. For example, if the ratio of pupils per school has a high variance from one year to the

[34]The exact phrasing of questions depended on the tier at which the respondent worked. Within each indicator, federal/regional level respondents are asked about the value of the 25th percentile district in their (chosen) region, the 50th percentile district in their (chosen) region, and the 75th percentile district in their (chosen) region.

next within the same district, this could be seen as a warning sign as to the quality of measurement. Table A2 presents summary statistics and tests for the presence of a random walk for each of the indicators used in the analysis where panel data are available. Column 4 shows the estimated autoregressive parameter from a Dickey Fuller Regression, controlling for district fixed effects and column 5 shows the estimated autoregressive parameter from a Dickey Fuller Regression, controlling for district fixed effects and district-specific linear time trends. The estimated coefficients are less than 1, which indicate stability in the trend, for all indicators except for three: the amount of land used for agricultural purposes; the amount of land used for pastoral purposes; and the number of business licenses issued. In addition, column 6 reports the z-statistics and column 7 the associated p-value from a unit-root test based on Harris and Tzavalis [1999] where there are sufficient data to do so. Column 8 reports the fixed-N critical value at the 5% level and column 9 the fixed-N test statistic from Im et al. [2003], allowing for linear time trends. These tests are appropriate for panel data settings with short timeframes. The results of the tests provide significant evidence against the existence of a unit-root process, and therefore a random walk, for the tested series. These results provide support to the notion that the benchmarking data are consistent over time.

We take the administrative and survey data as a proxy for the underlying state of the world. We use these variables to benchmark the claims of public officials. For each indicator, we subtract the corresponding value in the benchmark data from the relevant claims in our data set, such that a positive error is an overestimate. For example, the left-hand panel of Figure 1A provides histograms of both the benchmark data and public officials' claims regarding the population of the districts that they serve. The histograms are divergent at many points in the distribution, indicating that in general public officials are making errors about the underlying populations they are serving. The right-hand panel of Figure 1A provides a histogram and density plot of the scale of errors made by public officials in their estimates of the populations they serve.

The distribution of errors in Figure 1A is perhaps as would be expected. On this specific indicator, the average error is close to 0, and the distribution is close to being normally distributed. However, our data allow us to provide estimates of the scale of the standard deviation in the errors of public officials. The dashed vertical line represents the median population of a district in Ethiopia (120,398) to give a sense of the magnitude of the errors for this indicator. Many officials' claims are relatively accurate; 21.5 percent of public officials guess within 20% of the benchmark data. On the other hand, the scale of error of a substantial fraction of officials' claims is large; 16% of claims by officials over or underestimate the size of the claimed district by more than the population of the median district. 37% of officials overestimate or underestimate

the population of the district by more than half the population of the median district. And 39% of officials claim that their district's population is 50% bigger or smaller than it is.[35]

The scale of errors is similar across the distribution of district populations, implying that public officials are indeed adjusting their estimates towards the underlying state of nature. The left-hand panel of Figure 1B exhibits the scale of errors made by public officials in their claims about districts that have the 25th percentile smallest population in their region, the median, and the 75th percentile. The right-hand panel of Figure 1B exhibits the corresponding proportional errors. That is, the proportion of the benchmark population made up by the claim. In each case, over 50% of officials claim that they are serving a population that is at least 50% smaller or 50% larger than it is. Thus, across the full range of district sizes in Ethiopia, a sizeable fraction of public officials is making substantial errors as to the number of citizens they are serving.

Continuing to focus on the population data only, and restricting ourselves to the median-sized districts in each region, Figure 1C presents a series of thought experiments that allow us to assess the economic implications of the scale of the errors we observe. Suppose that federal and regional public officials were tasked to distribute public resources evenly across the district populations in their region of interest. The errors they make on the populations of individual districts are therefore compounded, as a district population the official underestimates doubly loses out to one she overestimates. The top panel of Figure 1C calculates the de facto weights the officials we study unwittingly embed in their distribution of resources across individual citizens by dividing the perceived population by the actual population. Each observation can be interpreted as the weight that a public official places on the median district in their region of interest relative to the weight that should be assigned when using the actual population data. A value of 1, therefore, signifies that the civil servants assigned a weight that accurately reflects the population data.[36] To the far right of the top panel we see that one official in our data would mistakenly weight the population of one district 3.5 times as high as the actual population would justify. And 14% of officials would weight the median district in their region at least 1.5 times more than its population justified. Relative to an equitable distribution of government resources across constituents, central officials

[35] A simple test of distributional knowledge is to ask federal and regional officials whether they can correctly rank the 25th, 50th and 75th percentile districts in their region of expertise in terms of population. In our data, only 18% of officials state the correct ordering for their region (this is true for both federal and regional officials separately). 47% of these central officials cannot accurately state which of the 25th and 75th percentile districts in their region has the larger population.

[36] Specifically, we calculate the weight assigned to each district in the region by dividing the population of the district by the total population of all districts in the region that were included in the survey. This is done using the claims from the civil servants and usng the actual population data. The relative weight is then the weight as implied by the claim of the civil servants divided by the weight as implied by the population data.

would over disburse by 30% to districts at the 75th percentile of the distribution of errors, and under disburse by 12% to those at the 25th percentile of errors.

The middle panel of Figure 1C restricts this analysis to managers only, on the assertion that managers may be critical in resource distribution decisions. The distribution is not dissimilar to that in the top panel, implying that more senior members of public service are vulnerable to mistargeting in a similar way to their more junior colleagues. The 75th percentile of managers and above in terms of overweighting would still distribute 1.5 times more resources to districts in their constituency than is justified by their actual population.

In the best case scenario, all members of the organization would have an equal vote in how resources were distributed so that measurement error would be minimized. Such a scenario is modeled in the bottom panel of Figure 1C, where the observations are now organization-averages. Here we see a reduction in the extreme tails of bias observed in the other two panels. However, a one-sample Kolmogorov-Smirnoff test for uniformity of the probability density function is rejected at standard levels of confidence, such that even capitalizing on the information of all members of an organization results in a bias towards some populations rather than others.

This reasoning implies that individual civil servants make economically meaningful errors on this simplest of public resource distribution tasks. Compounding this with other errors, such as the mis-measurement of primary pupil-teacher ratios, will distort the distribution of resources even further. Even in the best case scenario that capitalizes on all the information in an organization, such compounded errors would create substantial distortions in public sector choices.

The distribution of errors among other variables is similarly large. The estimates of 16% (23%) of public officials regarding the characteristics of their constituents are at least a standard deviation (half a standard deviation) away from the objective benchmark data we collect. The mean absolute error in education bureaucrats' estimates of primary enrollment numbers is 76% of the true enrollment figures. Across all of the education variables, the mean absolute error is 51% of the benchmark data. The mean error in estimates of the proportion of pregnant women who attended ANC4+ during the current pregnancy (the 'antenatal care rate') was 38% of the benchmark data. Agriculture officials overestimate the number of hectares in their district that are recorded as used for agricultural purposes by almost a factor of 2. As noted above, these magnitudes imply substantial misallocation of resources across space and subject matter. If bureaucrats think that antenatal care is a larger issue than it is, they will devise policies that prioritize corresponding investments beyond those that are justified by the underlying state of the citizenry.

Our core analysis aims to bring together the full set of claims in a unified framework.

However, individual errors are not comparable across indicators. Making an error of 0.5 in a district's pupil-class ratio is large for that indicator, but tiny for the population indicator. To jointly assess the scale of errors across multiple indicators, we undertook a number of transformations of the raw error data. First, we took absolute values of the errors so that deviations of different signs are comparable.[37] Second, we trimmed the data at the 95th percentile. In a very small number of cases, errors were so large that it seemed infeasible for even the most uninformed bureaucrat to have made such an error, and it looked likely to be a data entry problem. Either way, the interpretation of the remaining data is based on the errors of 95% of the claims of public officials in our study.

Our core transformation is to create z-scores of each of the variables by subtracting the mean and dividing by the standard deviation of all the errors related to a particular indicator. This scales the claims by an empirical measure of a 'normalized' measure of error for that indicator. The interpretation of the z-score is a unitless measure of how many standard deviations of the full distributions of error away from the benchmark data an official's assessment is. Thus, it is a measure of whether an individual official made a 'large' error relative to the empirical distribution of errors we observe for that indicator. Comparisons across indicators therefore take into account the empirical distribution of errors for each indicator. They compare the scale of an official's error relative to the errors made by all other officials in our data set, indicator by indicator.

Defining z-scores in this way assumes that the tails of the raw errors are roughly symmetric for each indicator, such that an error 10% below and above the truth are equally likely. However, the distribution of errors of a number of the indicators we study are right-skewed, such that we may be underestimating the severity of an error in the sense that few people would make it. To test the robustness of our core measure to this criticism, we define a 'skewness-adjusted' z-score, where we test for skewness of the distribution of a raw indicator and replace skewed indicators with their log. We then take the z-score of these transformed variables in a similar way to the above measure.

Finally, we move away from distributional concerns by taking the ordinal rank of absolute errors within an indicator and divide this by the maximum rank, thereby creating a 'rank-proportion' score. This can be interpreted as the percentile of error made by an official relative to other officials in our data set. The Data Appendix provides further details on how the error variables we use in the analysis were created.

Figure 2 provides cumulative distribution functions of these three measures. Figure 2A presents the distribution of the unadjusted z-scores, with separate plots for

[37]The probability of a negative (as opposed to a positive) error in the data we use for our core analysis is 0.404, such that the likelihood of over- and under-estimates is relatively equal. We have also transformed the errors to be percentages of the underlying benchmark data, and then created the various measures outlined below. Taking this approach does not qualitatively change our core results.

Ethiopia's three tiers of government. Since these represent z-scores of absolute errors, a lower z-score implies a more accurate claim. The figure highlights substantial variation in errors across government that echoes the variation in the raw indicators described above and we observe officials making the full distribution of claims. Previewing later results, we see that officials at the district tier of government make lower errors at almost every point in the distribution. A Kolmogorov-Smirnov test of distributional equality is rejected with a p-value of 0.00. Regional and federal organizations have almost identical distributions of errors, though regional organizations are slightly more prone to outliers.

Figure 2B plots the distribution of z-scores where adjustments are made to skewed underlying distributions when necessary. Conditional on differences in the ranges of the distribution produced by the two different methods, the graph is qualitatively similar to that of the unadjusted z-scores. The adjusted scores have a more even distribution of extreme values, as would be expected. Figure 2C plots the rank-proportion transformation, and while district errors continue to be substantially smaller than those at higher tiers, some distinctions arise between the regional and federal government levels. Overall, along the margin of decentralization across tiers, the three measurement methods provide highly consistent messages.

3.3.6 Measuring the Incentive Environment

To understand the incentive environments in which public officials in the organizations we are studying make decisions on information absorption, we also collected data on management practices. Following recent efforts to collect data on the management practices of public administrations Rasul and Rogger [2016]; Rasul et al. [2018*a*] we used a public sector version of the World Management Survey (WMS; Bloom and Van Reenen [2007*a*]; Bloom et al. [2012]) to elicit measures of the management practices under which public officials operate. Management interviews were 'double blind' so: (i) managers were not told in advance they were being scored or shown a score grid; (ii) enumerators were given no information on the organization's performance. Only managers (directors) were enumerated the World Management Survey and the measures of organizational management practices we use for our core analysis average management scores over the most senior divisional-bureaucrat reports where there are multiple directors.

The WMS evaluation tool elicits management practices through a semi-structured interview covering 7 topics: flexibility, incentives, monitoring, roles, staff involvement, staffing, and targeting. Table A3 details each of the 19 management-related questions, by topic, as well as the 1-5 scoring grid used by our enumerators for each question. To

provide a sense of the holistic (rather than specific) nature of these questions, we go through one example: a question relating to management practices relating to monitoring was, "In what kind of ways does your Directorate track how well it is delivering services? Can you give me an example?" Enumerators could then score responses on a continuous 1-5 scale, where for indication the scoring grid described a score of one as corresponding to circumstances where the "Directorate does not track performance"; a score of three corresponded to, "Directorate tracks a number of performance indicators. These are seen and reviewed by senior management only"; and a score of five corresponded to, "Full set of indicators are tracked formally and continuously. Reviews are conducted regularly and involve representatives of all directorate staff groups. The results of the review are formally communicated to all staff."

The scores on each question are converted into normalized z-scores (so are continuous variables with mean zero and variance one by construction) and are increasing in the commonly understood notion of 'better management'. For example, we assume greater monitoring corresponds to better management practice. We then take unweighted means of the underlying z-scores to aggregate the questions on monitoring shown in Table A3 and separately all other management practices. The 'Management Practice: Monitoring' index acts as a proxy for the organization's value of the information acquired by officials, and thus as a measure of U_{ii} and U_{ij} in AT.

We take two approaches to measuring alignment. The first, most consistent with our approach to measuring monitoring practices, is the creation of the 'Management Practice: Other' index, which aggregates the non-monitoring sub-topics of the WMS. These topics measure multiple aspects of management including staff involvement and incentives, all of which could be said to be means of aligning the preferences of agents to those of principals. For example, what are performance incentives but a mechanism for aligning agent action to the desires of principals? Appendix A3 details the questions included in this index.

The second measure complements this core approach by asking civil servants directly about the extent of alignment they feel to the organization and their daily activities. In the civil servants survey, we asked officials, "To what extent do you believe that the organization's mission is aligned to your own mission?" and "How consistent with your mission are the various tasks and activities assigned to you on a day-to-day basis?" To measure the extent to which the organization as a whole and the individual tasks assigned to the civil servant are consistent with her own mission, we create an alignment index that is the mean of the organization average response to the former of these questions and the proportion of employees in the organization that state "I am contributing to fulfilling that mission on an everyday basis" to the latter.

Figure A2 plots the distributions of the 'Management Practice: Monitoring' and

'Management Practice: Other' indices across the organizations we use in our core regressions.[38] Both indices exhibit substantial variation, indicating that de facto management practices vary substantially across the Ethiopian Public Service. For practices related to the organization's value of the information acquired by officials, the 75th percentile organization has a 'Management Practice: Monitoring' score that is 141% higher than the 25th percentile organization. For practices related to the alignment of preferences of agents within an organization, the 75th percentile organization has a 'Management Practices: Other' score that is 135% higher than the 25th percentile organization. this variation occurs despite organizations operating under the same broad political environment, the same federal administration, sharing the same national history, and working under the same governing rules over budget allocations, wages, non-wage benefits, promotion schedules and other contracting arrangements. This variation across organizations is consistent with evidence from other management surveys of the public sector [Bloom et al., 2015; Rasul and Rogger, 2016; Rasul et al., 2018*a*].

3.4 The Determinants of Superior Information in Hierarchy

3.4.1 Basic Descriptives on the Determinants of Information

Combining details on the characteristics of officials we surveyed and their organizations with the data on errors, we can provide broad descriptive evidence on the determinants of superior information in public hierarchies. Such descriptive evidence provides a broad framework in which to understand the later analysis.

Table A4 undertakes complementary ANOVA and OLS analyses where the dependent variable is the z-score of errors across variables outlined above. To understand the relative explanatory power of the variables highlighted by the literature in section 2, column 1 of Table A4 presents partial sum of squares for these variables with corresponding F-statistic p-values presented in parentheses. Column 2 provides coefficients for the corresponding variables within a formal regression framework for consistency with our later tables.

The first panel undertakes our first formal test of whether the hierarchical structure of government matters for the quality of officials' information along three margins. We see that tier of government is of substantial importance, and has a higher partial sum of squares than any of the other 'structure' variables. Its p-value is 0.00, which is consistent with AT's prediction that the structure of formal authority is a critical determinant of information absorption. In column 2, where the coefficients are from a

[38]The 'Monitoring', and 'Other' indices are positively correlated with each other, with a coefficient of 0.62. Importantly, the partial correlations of these two dimensions of management can still be separately estimated from each other.

regression of the variables listed in Table A4 on the z-score of errors, we see that the district government binary has a p-value of 0.02, previewing future results.

We also test the extent to which being of managerial rank impacts the scale of an official's errors, and again consistent with AT's predictions, the sum of squares is very low, and the associated p-value is 0.83. In environments in which the principal has formal authority (such as within an office), there is little incentive for agents to invest in information acquisition, equalizing the information absorption practices of principals and agents. The corresponding conditional coefficient in column 2 is 0.01, with a p-value of 0.86. We also find little evidence that there are substantial differences in errors across sectors, consistent with the idea that it is the structure of incentives, rather than the topic of focus, that matters for information acquisition.[39]

The marginal costs of information acquisition may be determined by features of the variables we test officials on, the characteristics of the responding official, their organization's infrastructure for disseminating information, and features of the local environment. The second panel of Table A4 outlines a corresponding series of variables that increase or decrease the marginal costs of information acquisition. First we focus on features of the claims, and see that the specific indicator on which officials make claims explains a significant portion of the variation in overall error rates, as one might expect. The percentile of an indicator (whether it is the 25th, 50th or 75th percentile district in the data) similarly impacts the scale of error, with the average being a more familiar quantity to officials than other points in the distribution. These results imply the importance of controlling for indicator features in our core regressions.

Being a more experienced or educated bureaucrat may lower the cost of information acquisition. However, the coefficients in column 1 corresponding to bureaucratic characteristics indicate that individual demographics have little impact on the officials' overall likelihood of making errors. The corresponding regression coefficients are all insignificant at the usual levels. The sum of squares is higher for organizational features, and the results are consistent with the predictions of AT. For example, AT argues that "it is always optimal for the firm to be in a situation of overload" where managers credibly commit to lower monitoring of employees through increased span of control. A higher span of control therefore leads to lower errors on average among officials, as employees increase their investment in information. We see that the corresponding coefficient on span of control in column 2 of Table A4 is negative, implying that operating under a higher span of control leads to officials making lower errors in their claims.

[39]Though not shown in the table, we have also investigated the impact of environmental features on the scale of errors made. These are lower in districts with a lower poverty rate, smaller population, and those that are less remote. However, these features explain only 1.3% of the variation we observe, and once again highlight the primacy of incentives in determining the quality of information a public official has.

To ensure consistency with later analysis, we include district fixed effects in Table A4. To explore the notion that simple proximity to local constituents (such as being an official at a local-level organization) is driving differences in information quality, we can drop these fixed effects and include measures of 'remoteness' of the district: the average travel time of constituents to the nearest urban area, the average elevation of the district, the average slope of the district, and the road density of the district. The partial sum of squares of all of these variables combined is 2.76. While non-trivial (and accounted for in our main specifications with district government fixed effects), these variables combined account for less variation in errors than variables capturing management practices or decentralization. In the OLS regression similar to column 2 of table A4, removing district fixed effects and including these variables, none are statistically significant at usual levels and a joint Wald test of their significance gives a p-value of 0.546. These results are consistent with the argument put forward in Cremer et al. (1994) that simple proximity is unlikely to be a key driver in differences in information across tiers of government.

The final panel of Table A4 assesses factors related to the marginal benefits of information acquisition. We split our measures of management practices into those related to monitoring and all others. Management practices related to monitoring explain the largest portion of the variance in errors. Previewing future results, the coefficient on monitoring-related management practices in column 2 of Table A4 is negative, as we would expect, and significant at the 5% level.

Alignment between the mission of an organization and its employees is a feature of the principal-agent relationship frequently discussed in the literature. Table A4 includes our complementary measures of alignment. The measure based on the management practices approach explains substantially more variation than that based on direct questioning. However, the coefficients on neither variable are significant at the usual levels in column 2 with p-values of 0.138 and 0.431 respectively.

The descriptive evidence laid out in Table A4 is consistent with the notion that formal power, and mediating management practices related to information use, are the key drivers of information acquisition in public organizations. The nature of incentives, rather than the individual characteristics of officials or the sector in which they work, explains the dominant portion of variation we observe in the information individuals hold. We now go on to add greater structure to our analysis of the specific predictions in the related literature.

3.4.2 Prediction 1: Formal Authority is Fundamental

Theoretically, principal-agent relationships in which the principal has formal authority over the agent limit the incentive for the agent to invest in information acquisition. In the Ethiopian government, this implies that employees have no greater incentive than their managers to acquire information. It is then an empirical question as to what the residual difference in knowledge of the local environment will be between managers and employees. Similarly, relationships in which the agent has formal authority should be characterized by their superior information on local conditions. In the Ethiopian government, this would imply claims by district officials will be more accurate than those by central government officials.

Our main estimates of the impact of hierarchy on information uses an OLS regression of the following form,

$$y_c = \gamma_1 \mathbf{1}\{\textbf{District} = \mathbf{1}\} + \gamma_2 \mathbf{1}\{\textbf{Manager} = \mathbf{1}\} + X_c + \epsilon_c \tag{17}$$

where y_c is a dependent variable relating to the scale of the error of the claim, such as the z-score of the absolute error, $\mathbf{1}\{\textbf{District} = \mathbf{1}\}$ is a dummy that takes the value 1 if the claim is made by a district official, $\mathbf{1}\{\textbf{Manager} = \mathbf{1}\}$ is a dummy that takes the value 1 if the officer making the claim is a manager, and X_c is a set of sector, district, indicator and percentile fixed effects, individual control variables, marginal cost and noise controls. Individual-level controls are tenure in the civil service of the respondent making the claim; tenure in position; tenure in organization; an indicator for whether the respondent is female; an indicator for whether the respondent holds an undergraduate degree; and, an indicator for whether the respondent holds a master's degree. Given the proximity of the federal and regional errors, as described above, we aggregate these into a single 'centralized' tier. When we run the regressions with all three tiers of government, the federal coefficient is not significantly different to the regional at the usual levels, as will be documented below.[40]

The regressions are at the claim level, so comparisons can be made across individual percentile-indicator pairings, providing the most precise comparison we can offer. In our preferred specification, the variation that identifies the core coefficients is that between two sets of claims, made about the same indicator, at the same percentile of the corresponding distribution, in the same sector, about the same district by individuals at the two levels of hierarchy, conditional on their individual characteristics. Thus, our specification ensures a relatively tight comparison between levels of the hierarchy. We

[40]The fact that there is little difference between the federal and regional tiers is also evidence against physical distance or remoteness playing a critical role in the acquisition of information in this setting [Agarwal and Hauswald, 2010*a*; Landier et al., 2007].

restrict our sample to those claims that are made about the same districts to ensure the variation in our regressions arises only from coherent comparisons. That means the claims used for comparison are made by 594 federal and regional government officials, and 297 district officials. Since there may be correlations in the sectoral information about a particular district, we cluster at this level throughout our analysis.

Table 2 presents the results. In columns 1 and 2, the dependent variable is a binary that takes the value 1 if the individual official stated that they did not know the answer to a question, and thus could not make a claim. We see in column 1 that district officials are 17 percentage points less likely to answer 'Don't Know', while there is almost no difference between managers and non-managers in readiness to provide an estimate. Controlling for sector, district, indicator, percentile and individual controls in column 2 of Table 2 does not change these results. This implies that if you were to enter offices at different tiers of government, you would face a richer information environment at the district level even conditional on the level of errors. The error results should therefore be interpreted as 'the perceived state of knowledge within an office', and imply they represent a lower bound on the difference in knowledge between the two tiers.

Columns 3 to 5 of Table 2 present unconditional versions of (17) where the dependent variable is the z-score of errors across socio-economic variables. In column 3 we include a district government dummy only. We see that across all claims those made by district officials have significantly lower errors than those of central officials. In contrast, column 4 includes only a manager dummy and indicates that there is no difference in the magnitude of errors made by managers and their employees. In column 5 we include both dummies, and find that the results do not change.

We can also assess these unconditional results graphically. Figure 3 presents the average errors of centralized and district officials in their claims about each of the 42 districts on which we focus our core analysis. The set of claims over which these averages are made are the same as in our regressions such that they are balanced across variables and percentiles. The scatterplot is ranked by the average error of claims made about each district, and we can immediately see that some districts are fundamentally more difficult to make assessments about. Corresponding to the results in Table 2, we see that the unconditional means of claims by district officials are typically below those of officials in centralized organizations across the distribution. In 75% of cases, the district officials' assessments are more accurate than those of the corresponding central government officials.

This pattern of results is reflected in regressions using variation from within the sector (column 6), constraining comparisons to be based on claims about a common district (column 7), focused on common indicators (column 8) and on common percentiles of those indicators (column 9). Column 10 presents our preferred specification,

where we also condition on characteristics of the individual official making the claim. It presents some of the most precise empirical evidence we know confirming a classic assumption in organizational theory; that agents have superior information about local conditions. However, in these data we find that agents with formal authority (those acting in district organizations) have superior information, while those who do not (those acting under a principal within the same office) have no better information than their seniors.[41] The superior information of agents is a product of a particular form of decentralization rather than a rationale for it. In terms of the geography of information in the public sector, these results imply that information is spread relatively evenly across layers of the hierarchy within an office, but that significant asymmetries occur across offices.[42]

The coefficient on the district government binary is -0.33, implying that claims made by district officials are on average 0.33 standard deviations closer to the benchmark data than central governments across the z-scores of variables related to citizen characteristics. As can be seen by comparison to the coefficients in Table A4, such institutional features dwarf the impact of greater tenure in the service or increased formal education. The coefficient on district government is also an order of magnitude more important than if an MIS system exists within the corresponding organization. If we remove district fixed effects to allow for estimation of the impact of district poverty rates on errors, we see that the coefficient on the poverty rate (measured from 0 to 100) is 0.008 with a p-value of 0.05. Thus, officials make greater errors in poorer districts, as we might expect. A one percentage-point increase in poverty is associated with a 0.008 standard deviation increase in the absolute error. The size of the coefficient on district government in Table 2 is 41 times larger in absolute terms, such that the quality of information in district organizations is akin to estimating over the 25th percentile of poor districts rather than the 66th.

To assess the robustness of these core results to our methods of measurement, we present a series of robustness checks in Table A5. The first column of Table A5 re-runs our preferred regression, but disaggregating the central tier of government binary into the federal and regional tiers, with region as the excluded category. The coefficient on district government is -0.31 with a p-value of 0.00, equivalent to that in our preferred regression, and the federal government binary has a value of 0.04 and is insignificant

[41] In our preferred specification, the interaction between district government and manager is not significant at the usual levels, with a coefficient of -0.07 and a p-value of 0.58 (the coefficient on district government is -0.307 with a p-value of 0.004 and the coefficient on manager is 0.053 with a p-value of 0.52). Thus, at both central and district tiers of government, there is a common information environment within the organization.

[42] This assertion is supported from statements by officials in our survey. Only 4% of central level officials (1% of federal and 6.5% of regional) select 'formal interactions with lower-tier institutions' as their top source of information.

at the usual levels, implying that there is no statistical difference between the errors made by federal and regional officials. As described above, the federal and regional organization structure is equivalent and functionally distinct to the district. In columns 2 to 5 of Table A5, we return to a binary district variable and vary the level at which we cluster the standard errors in our preferred regression. Taking the conservative perspective that information regarding a specific district may be freely shared among officials across government, we cluster at the 'claimed about' district in column 2. The district coefficient continues to be significant at the 1% level. Similarly, we may be concerned that errors are correlated across the respondent's claims, within their organization, or across the region they work in through their regional sector offices (who provide them with regulations and technical assistance). We therefore cluster at the respondent level, the organization-of-respondent level, and the region-sector level in columns 3, 4 and 5 of Table A5 respectively. In each case, the coefficient on the district government binary is significant at the 1% level.

We can also rerun our preferred specification using proportional rank as an alternative measure of the magnitude of errors. Column 6 of Table A5 presents the result, which implies a district official makes an error that is 10 percentage points lower than a central bureaucrat. Returning to comparison with the poverty rate, this magnitude is equivalent to 37 times the magnitude of the coefficient on the poverty rate using this dependent variable, and thus akin to making claims over the 25th percentile of poor districts rather than the 62nd. As outlined in section 3.3.5, we may be concerned about using absolute errors in the presence of skewness of the underlying indicators. We therefore re-run our preferred specification on the 'skewness-adjusted z-score' and find that the coefficient on district government is even larger in magnitude with a coefficient of -0.39, and a p-value of 0.00. Distinct measurement approaches produce qualitatively similar results in this setting.

As indicated in Figure 3, the mean error varies substantially across claimed about districts. It seems natural to hypothesize that agents would have a greater information advantage under de facto agent-authority in environments where information is costlier to acquire. However, totally differentiating the utility functions in AT shows that increases in costs to both agents typically have an ambiguous effect on the relative distribution of information. For example, differences in the shape of the cost function at the specific levels of investment will play a role in determining the effect on the relative distribution. It is therefore an empirical question as to whether the differential between principal and agent increases as the contracting environment becomes more difficult. To assess this in our setting, Table A6 presents the results of a quantile regression of our preferred specification over the quantiles of absolute error. The coefficient on the district dummy is negative and significant at the 1% level across all percentiles of

absolute error, implying that district officials' claims are closer to the benchmark data than centralized officials' claims at all levels of absolute error. There is no clear pattern across the percentiles in terms of the size of coefficients, and 6 of the 9 are between -0.20 and -0.27. Thus, the impact of formal authority in this setting is not mediated by the scale of the error to which it relates, but rather has broadly positive impacts on information acquisition.[43]

3.4.3 Prediction 2: Management Matters

The diversity in management practices across Ethiopia's public organizations, and specifically in terms of how highly information is valued as an organizational asset, can be interpreted as variation in the rewards to officials of acquiring information. Organizations that value the use of information, captured by the 'Management Practices: Monitoring' index, create a high relative reward for obtaining information. In AT, the responsiveness of actors to the introduction of a de facto agent-authority regime is mediated by the value of holding that information. In environments in which organizations value information, captured in the 'Monitoring' index, we should expect to see lower errors. Conversely, AT highlights that management practices that raise the alignment of the principal and agents' interests may lead to a strategic substitution of effort in acquiring information.

We can extend our analysis in (17) to include measures of management practices and assess the consistency of these ideas with our data. To do so, we broaden the variation across which we make comparisons by switching from district fixed effects to regional fixed effects. Our reasoning is that since management practices vary at the organization-level, within a district we constrain the variation to be across 5 organizations. By extending our estimating variation to the region, we are able to capitalize on variation across an average of 122 organizations while still constraining the comparisons

[43]Tacit knowledge of local conditions can be characterized as 'soft information', which cannot be verified by parties other than those who collected it. It could be argued that the Ethiopian government's 'Public Service Rules', written in hardcopy for all officials, can be characterized as 'hard information'. Given its verifiability, Dessein [2002] argues that soft information leads to inefficient centralized decision making stemming from noisy communication, while communication is preferable to delegation if there is little uncertainty so that headquarters could make decisions primarily based on hard or public information. Table A7 presents a form of equation (17) in which the dependent variable is the z-score of the absolute error of claims about the public service rules. We find that across our specifications the Federal Government makes lower errors over hard information than either of the other two tiers of government. Column 7 is equivalent to our preferred specification in Table 2, and it implies Federal officials make errors over the public service rules that are 0.09 of a standard deviation lower than regional colleagues, who make errors that are statistically insignificantly different to those at the district level, at the usual levels of significance. This result is robust to clustering at the respondent level (column 8 of Table A7) and the region-sector level (column 9), and to using the skewness adjusted z-score (column 11). Column 10 indicates that the difference is equivalent to a 4 percentage point reduction in errors.

of errors within a restricted geographic and institutional environment. We also follow standard practice in studies of the World Management Survey indices, and include 'noise controls', which aim to soak up aspects of the enumeration of the management survey. Noise controls are composed of the time of day of the survey, day-of-survey fixed effects, enumerator fixed effects, an indicator for the enumerator's subjective assessment of the quality of the interview, an indicator for the duration of the interview (decile fixed effects), controls for the average tenure and gender of the managers interviewed. Finally, to focus the analysis on the benefits of acquiring information, we condition on proxies of the marginal cost of that acquisition. Marginal cost variables are the proportion of managers in the organization that state that a management information system is in place; the organization average response to "In what proportion of projects would you say information flows effectively between individuals in the same organization?"; the organization average response to "In what proportion of projects would you say information flows effectively between organizations?"; the organization average response of managers to "How many personnel do you manage?"; the organization average response of employees to "How many people would you say regularly give you tasks as part of your formal work duties?"; and, the additional alignment index that complements our main measure. To ensure the results of Table 2 tolerate these changes, we present our preferred specification with these modifications in column 8 of Table A5. We see that the coefficient is -0.28, still significant at the usual levels though now at the 10% level.

Table 3 presents the results. In Column 1, we include the management indices, the manager dummy, and only those controls equivalent to our preferred specification in Table 2. We see that the coefficients on the management indices accord with theory exactly. Increasing the value an organization places on the use of information, proxied by the 'Management Practices: Monitoring' index, decreases errors substantially. The corresponding coefficient is -0.17, and the p-value 0.014. Similarly, the coefficient on the 'Management Practices: Other' index is 0.28, and has a p-value of 0.005. Managers continue to have information equivalent to their employees. These results are slightly strengthened by the inclusion of noise controls and proxies of the marginal cost of information acquisition, as described in column 2. Combining the indicator of de facto authority, the dummy for district government, and the management practices that proxy incentives leaves the qualitative results consistent with the literature. Once again, these results indicate the primacy of incentives for determining the distribution of information in the public sector.

In column 4 of Table 3, we undertake a first assessment of whether the impact of de facto authority varies by the nature of management practices. We include an interaction between the monitoring management practice index and the district government binary. It is negative, and large in magnitude relative to the baseline impact of de

facto authority, but not significant at the usual levels, with a p-value of 0.17. However, we take this heterogeneity analysis further in Figure 4 by plotting the distribution of coefficients of the district government binary as we vary the decile of management practice. The figure makes clearer that the impact of de facto authority is clearly mediated by incentives. At the lowest levels of management practices observed in our setting, where incentives to acquire information are weakest, providing agents with de facto authority has zero impact on their relative magnitude of errors. At the highest level of management practices, the corresponding coefficient is -0.66 with a p-value of 0.04. These results imply that unconditional decentralization, without effective corresponding management practices or more broadly incentives for information acquisition, does not improve the information environment. This finding could help reconcile the divergent findings on the impact of decentralization on service delivery (documented in surveys such as Bardhan [2002] or Crawford and Hartmann [2008]), as management practices are typically not measured in such studies.[44]

Returning to Table 3, we assess whether management practices have a differential effect by whether an official is a manager. Organizational monitoring practices have slightly stronger effects for managers, but the coefficient for non-managers is still negative and significant at the 5% level, while the coefficient relating to the interaction is insignificant at the usual levels.

We can also assess the baseline impacts of management practices that align principals and agents. Since our management practices are measured at the organizational level, this implies we are in the realm of de facto principal authority, in which AT states that greater alignment should lead to decreases (increases) in principal (agent) information acquisition. However, as seen in section 3.2, the predictions of the wider literature relating to alignment vary depending on the precise institutional environment. The dominant effect of alignment in our setting, as measured by the 'Management Practices: Other' variable, is to increase errors across the columns in Table 3. The coefficient in column 3 is 0.32 and significant at the 1% level, thus large compared with our other coefficients of interest.[45] We do not find any evidence of these results varying across the hierarchy. The coefficient on an interaction of the manager binary and the

[44]In the setting of our study, we do find evidence that management practices vary across tiers. Table A8 reports on organization-level regressions where our management indices are the dependent variables, the district government binary is the explanatory variable of interest, and a range of controls are presented. We find that once differences in individual characteristics and sector fixed effects are conditioned on, the coefficient on district government is large (0.41) and significant at the 5% level, indicating that on average, district governments have more of a culture of monitoring. This finding is strengthened once we condition on differences in the marginal cost of information acquisition. We also find weaker evidence for other management practices.

[45]We do not find equivalent baseline effects using our secondary measure of alignment, which directly asks officials how aligned they feel. The corresponding coefficient is 0 and insignificant at the usual levels.

'Other' index is 0.00, with a p-value of 0.97. Our results therefore imply that increases in alignment have a broadly negative impact on the incentive to acquire information as officials aim to free-ride on their colleagues efforts.

3.4.4 Information, Promotion, and Service Delivery

Before turning to our experimental evidence, we are able to provide descriptive evidence on the impact of information acquisition on an individual official's career trajectory. By regressing the data we have collected on individual appraisal scores, we can assess how the magnitude of errors an individual is likely to make in their claims about citizen characteristics correlate with their career progress. To ensure that we take account of differences in scoring systems across organizations, we convert appraisal scores into organization-specific z-scores, and cluster at the district government-sector level. Since our interest is beyond comparisons of common claims across the hierarchy, we expand our sample in this regression to all claims at organizations at which we have appraisal data, leading to an expansion in the number of claims we estimate over to 3,963 claims at 135 organizations.

Table A9 presents the results. Column 1 indicates that unconditionally, a standard deviation increase in absolute error is associated with a reduction in appraisal scores of 0.02 of a standard deviation. Controlling for tier (column 2), sector (column 3), indicator (column 4), percentile (column 5), and region (column 6) leads to a coefficient of -0.028, which is significant at the 5% level. This is true if one controls for individual controls (column 7), though the characteristics of officials account for some of the variation, and for differences in organizational characteristics (columns 8 and 9).[46] Thus, information acquisition is positively correlated with incentives in the public sector, consistent with the results in the previous section. Interestingly, admitting to not knowing the characteristics of the citizenry does not have significant impacts on an official's appraisal scores in any of the specifications presented.

Beyond career concerns, information may have impacts on the quality of public services delivered by the bureaucrats we study. We therefore use z-scores of the underlying service delivery indicators that we use to build the error indicators as a dependent variable. Table A10 presents the results of a regression of absolute error (variously measured) on these outcomes. Again, since our interest is beyond comparisons of common claims to a focus on service delivery, and to ensure some degree of homogeneity in incentive environment, we expand our sample in this regression to all claims at district organizations for which we have corresponding service delivery data. This leads to

[46] In terms of the impact of basic characteristics on appraisal scores, managers perform significantly better on the evaluations, while females achieve signficantly lower scores at the usual levels. Education and tenure are not significantly associated with the evaluations at the usual levels.

expansion in the number of claims we estimate over to 2,853 claims at 276 organizations.

Throughout the specifications in Table A10, the absolute error does not have clear impacts on service delivery outcomes. Once sufficient controls have been conditioned on, officials stating that they 'Don't Know' has a negative impact on service delivery outcomes, of the order of 0.2 standard deviations of the average service delivery indicator z-score. This is akin to an education official stating 'Don't Know' leading to a reduction in the enrollment rate of 9.2 percentage points, or in health of a reduction in the antenatal care rate of 6 percentage points. More research is needed on the precise impact of bureaucratic information on service delivery outcomes. In our setting, Somani [2018] finds evidence that the historical implementation of the education management information system increased enrollment rates by 2.7 percentage points. Similarly, other studies of improvements in bureaucratic information have tended to find positive effects on service delivery (for example, on public health worker attendance in Callen et al. [2018] and Dhaliwal and Hanna [2017], and on the implementation of India's rural employment guarantee in Muralidharan et al. [2016]).

3.5 Experimental Evidence on Information in Hierarchy

3.5.1 Experimental Design for Investigating Prediction 3

We capitalized on the relative feasibility of an experiment to test the third core prediction of section 3.2. By dramatically reducing the cost of information acquisition for all officials, we override the free-riding concerns inherent in the systemic nature of information acquisition in teams. AT assume stability conditions that yield increases in information from all reductions in cost functions. However, since such conditions are empirically unverified, we aimed to design an intervention that would *symmetrically* reduce principal and agent costs *dramatically*. This ensures the validity of our experiment under a wider range of conditions than AT assume. Even in the presence of free-riding concerns, the symmetric reduction in costs of information acquisition our intervention represents overrides these concerns, and we gain a clear prediction that information acquisition should increase. See the Appendix for details.

Prior to the enumeration of the survey, a random sub-sample of the organizations we study were sent the official administrative data corresponding to our policy variables of interest as a government circular. A set of federal organizations received data for all regions for all indicators, highlighting the median district for each indicator for each region. Similarly, a random sub-sample of district jurisdictions were sent the official administrative data for their whole region for all indicators, also highlighting the median district for each indicator. Since the information package to the district-level jurisdictions had to pass through the regional offices, this meant that all regions were

treated.[47] Since all regions were asked to communicate with the district administrations of the arrival of the survey team in advance (to prepare the logistics, in terms of office space, in terms of timing etc), this rules out concerns that the letter itself, and not the information within the letter, is driving the treatment effects of the experiment.

The information package was designed to accord exactly with standard communications of the Ethiopian Public Service. The package, and accompanying letter, was designed in close collaboration with the Ministry of Public Service and Human Development, and was sent through the standard bureaucratic channels. Thus, it was representative of broader service communications, and as far as we understand it would have been received as such. The accompanying letter stated that, "we have assembled the most up-to-date administrative data available to provide an understanding of the context your office works within. We would like to provide you with this information as an input to improved service delivery within your organization." The packages were sent along with communications about the survey, from May 2016 onwards. The first surveys were conducted in mid-June and lasted until mid-September. Letters and data packages were typically sent two weeks before the survey. The package was shared as a circular, and so the expectation was for it to be shared widely among staff within a district office.

The data packages make information on core variables of national policy available to a random subset of Ethiopian officials in a form they are used to at close to zero cost. The package is a clear reduction in $g'(e_i)$ from section (3.2). Officials must then decide to read and absorb the information, and make that decision in the wider incentive environment we have documented above.

3.5.2 Prediction 3: Low Cost Information Has Effects

Table 4 presents the results of the intervention on the errors made by officials in their claims about the socio-economic characteristics of the citizenry they serve. The regression framework has the same structure as (17), the sample is as in previous tables, and the dependent variable is once again a function of the absolute errors made by officials. We include a binary treatment variable that takes the value 1 if the organization received the data package, and 0 if it did not.

[47] Randomization was undertaken using random.org. For the treated district jurisdictions, the information package was sent to the Regional Bureau of Public Service with instructions to pass on the data to the randomly selected districts. This primarily occurred through the Woreda Office of Public Service, which then disseminated the information to the sector offices within the district jurisdiction. Table A11 presents a comparison of treatment and control groups across key characteristics of individual officials and the organizations in which they work. Treatment and control look balanced across individual and organizational characteristics. Treatment and control groups differ on a single variable at the usual levels of significance (years in position).

Column 1 of Table 4 presents the unrestricted treatment effect, which does not restrict the comparison of claims to be within indicator, percentile, district, and so on. We see that it is positive and statistically insignificantly different from zero at the usual levels. Throughout Table 4, we present p-values of the treatment effect with two distinct approaches to clustering. In the square and curly brackets below each treatment coefficient we follow the rest of the paper in clustering at the district government-sector level. These clusters correspond to the coherent sector teams within the district-government in which many information distribution choices would most likely be made. However, we note that the treatment is at the district level, and therefore we also present the p-value of the treatment coefficient from the same regression but clustered at the district level. In neither case is the fully unconditional treatment effect in column 1 significant at the usual levels.

In column 2 of Table 4, we condition on the set of controls used in previous tables, where the treatment effect relates to a comparison within indicator, percentile, sector and region, with individual and organizational controls. The impact of the data package is now negative and significant at the 10% level (a p-value of 0.08 using our standard clustering approach; a p-value of 0.06 clustering at the level of treatment). Its magnitude is roughly half of the coefficient on the district government dummy in Table 2, and large compared to the control variables used for comparisons above. It seems that officials have taken the opportunity of the reduction in the cost of information acquisition and increased their knowledge of the constituencies in which they work.

Given that we have documented the importance of de facto authority, we control for our core explanatory variables from Table 2 in column 3. The treatment effect jumps once we control for these critical drivers of variation in claim error. The coefficient, -0.28, has a p-value of 0.004 when using our standard clustering approach and 0.003 when clustering at the treatment level. Once we control for treatment, the impact of a claim being made by a district official increases to -0.44, and the coefficient on manager continues to be small, which is evidence that the experiment is not driving our results in Tables 2 and 3. A final set of controls is the addition of the management indices assessed in Table 3, 'Management Practices: Monitoring' and 'Management Practices: Other'. The treatment effect continues to be large, with a coefficient of -0.23, and corresponding p-values of 0.04 with our standard clustering approach and 0.03 when clustering at the treatment level. The management indices both have the expected signs and are significant at the 5% level. Thus, controlling for the key drivers of information acquisition, we see that a substantial reduction in transaction costs can improve bureaucratic knowledge.

The treatment effect is robust to using the different measurement techniques for aggregating absolute error. Using proportional rank of error as the dependent variable,

we find a treatment effect of -0.074, with corresponding p-values of 0.03. The effect weakens slightly when we use the skewness-adjusted z-score, but is still significant at the usual levels. Finally, we note that since the data package was sent by sectoral officials at the regional level, we should also test the robustness of the effect to any correlations between activities within the corresponding regional office. We therefore cluster at the region-sector level and the p-value raises to 0.048.

Together, the results support the prediction that substantial reductions in the cost of acquisition for officials can improve the accuracy of their information. This could be related to external efforts to improve the knowledge of public policymakers by presenting key findings of research in digestible, 'low-cost' ways, and by 'translating' research findings into a medium of communication that reduces the transaction costs to public servants of absorbing relevant messages. The magnitude of the coefficients in Table 4 are large relative to the incentive impacts documented in Tables 2 and 3, and substantially larger than many contextual factors, such as making claims over higher poverty districts. Providing a data pack to an organization improves the accuracy of its officials' claims by a similar magnitude to if they were estimating the conditions of the 5th percentile of poor districts rather than the 90th percentile.

However, as we have seen in previous tables, the conditions under which public servants work are highly variable, and estimates relevant to the 'average' bureaucrat may not provide a full picture of the impacts of our treatment. In Table 5, we therefore assess the heterogeneity of the treatment effect along the critical margins for information acquisition outlined throughout the paper. Column 0 presents our baseline treatment effect from Table 4 for comparison.

Column 1 presents the baseline specification in column 0 with the addition of an interaction between the treatment binary and the district government binary. The coefficient is not significant at the usual levels, with a corresponding p-value of 0.14. Similarly, column 2 presents the baseline specification with the addition of an interaction between the manager and treatment indicators. The interaction is once again not significant at the usual levels, and the p-value 0.37. There is thus no evidence that treatment varied substantially by the level of authority. As expected, in both authority regimes, a large decrease in the cost of information acquisition increases the equilibrium level of knowledge.[48]

Where we should expect to find results is in the incentives associated with information. In equilibrium, agents weigh up the marginal benefit of acquisition with the cost. Where benefits are low, small changes in the cost of acquisition can have large impacts on the equilibrium calculus for an agent, and our treatment should have large

[48]Interestingly, there is no evidence of a shift in the distribution of information between managers and employees. This is consistent with the treatment having been applied consistently within offices.

effects. Where the benefits of holding knowledge are high, information holdings will already be substantial, and thus small changes in costs will lead to limited additional information absorption. We test these ideas in column 3 of Table 5, where we interact our treatment effect with the 'Management Practices: Monitoring' index. The results closely match the theoretical prediction. At the mean level of management, we observe a similar treatment effect to the baseline. However, there is a large positive interaction effect, with a coefficient of 0.34, significant at the 1% level. As management practices related to monitoring improve, the mean treatment effect is attenuated. This can be seen clearly in Figure 5A, where we plot the treatment effect at the deciles of the monitoring management practice index based on the specification of column 3 in Table 5. At the lowest levels of management observed in our data, the treatment effect is extremely large. The coefficient is -0.87 (p-value 0.00), which is almost three times the average effect of a claim being made by a district official. However, from the 60th percentile of management practices related to monitoring upwards, the treatment effect is indistinguishable from zero at the 5% level. The best managed organizations incentivize their officials such that they already have sufficiently precise information that our intervention has no effect.

In column 4 of Table 5, we interact the treatment binary with the 'Management Practices: Other' index. Once again, the interaction is positive and significant at the 1% level. The treatment effect at the mean level of the management index is negative, implying that once again, a higher score on the management index attenuates our treatment effect towards zero. This can be clearly seen in Figure 5B, in which we plot the treatment effect at the deciles of the 'Other' management index, with all the controls of the specification in column 4 of Table 5. The figure presents a distribution of coefficients similar to that for monitoring, with substantial treatment effects at low levels of management, and a null effect at the 60th percentile onwards. Once again, this is as theory would predict. At low levels of staff alignment, agents will be more responsive to reductions in costs of information acquisition as there is little crowding out from free-riding incentives. However, as the organization implements management practices that align actors' incentives more closely, the incentive to absorb information lessens and agents aim to free-ride on each other's efforts. Overcoming free-riding concerns with a reduction in acquisition costs becomes harder.

Together, these interactions are consistent with the idea that organizational incentives are important mediators of the data package treatment. Where incentives for information absorption are strong, officials will source information relevant for their core work without external intervention. Where incentives are weak, there is room for external intervention to improve bureaucratic knowledge. In these poorly managed environments, external provision of low-cost information can have impacts on the

magnitude of officials' errors equivalent to or greater than the design of incentives. However, improving the underlying incentives for information acquisition would yield superior information outcomes without external intervention.

3.6 Conclusions and Further Research

This paper investigates the determinants of the distribution of information across the hierarchy of a public bureaucracy, the Government of Ethiopia. It measures the subjective beliefs of public officials as to the characteristics of the citizens that they serve at all three tiers of government and across five major sectors. By comparing these beliefs with benchmark objective data from surveys and administrative data, we identify which officials have the most accurate perceptions of the characteristics of their constituents. Evidence from a survey of these officials implies that such 'tacit' knowledge is the primary source of information for operational decision making.

Overall, the descriptive and experimental evidence is clearly consistent with the theoretical literature laid out in section (3.2). The distribution of information across bureaucracies is dominantly determined by the incentives of agents to undertake costly investments in information acquisition and absorption. These considerations dwarf impacts of the characteristics of the information being considered. Under the appropriate incentives, officials can hold relatively accurate beliefs regarding their constituents. Such incentives include holding de facto decision-making authority, working within an organization whose management practices value officials' holding of information, and being able to access information at very low cost. However, without these incentives, officials can make very large errors about the constituents they serve - roughly half of those bureaucrats we study cannot correctly rank the 25th and 75th percentile of districts in terms of population in the region in which they specialize. Such errors can have large economic consequences as they skew the distribution of resources. Relative to an equitable distribution of government resources across constituents, central officials would over disburse by 30% to districts at the 75th percentile of the distribution of errors, and under disburse by 12% to those at the 25th percentile of errors.

Given the importance of tacit knowledge in bureaucratic settings, the paper presents some of the most detailed evidence on the nature of information in the public sector we know of. It provides empirical evidence on the impacts of decentralization on the beliefs of individual public officials, on the role of management practices as mediating factors in the impacts of authority, and on the role of external information provision in a non-market hierarchy. While the empirical setting of this paper is the public sector, it is reasonable to assume that the same results would be observed where the same

hierarchical conditions exist in the private sector.[49]

The paper provides a framework for understanding bureaucratic constraints to knowledge diffusion in the public sector. While Brynjolfsson and McElheran [2016] document the rapid adoption of data-driven decision making in the private sector in the United States, World Bank [2016*b*] documents the slow rate of adoption of similar projects in public sectors across the world. Much policy-focused research aims to translate its findings into a form digestible by public officials. However, our results imply that embedding that knowledge into public organizations is a function of the de facto authority and corresponding incentive environment in which a public official works, as much as it is a function of limiting the costs of absorption over which the researcher may have some influence. The open question is how policy-focused researchers should respond to the incentive environments they face across the public sector.

A remaining question is the extent to which the precision of information within bureaucracy has impacts on service delivery outcomes. The objective of this paper was to understand the geography of information in hierarchy. Further evidence on the nature of the public sector production function would enable us to understand the relationship between the quality of information held by public officials, and the quality of the corresponding public services. Where accurate information is found to have substantial impacts on service delivery outcomes, it would be useful to understand how management practices could be designed to maximize service quality. Recent evidence implies there may be a tradeoff between generating incentives for officials to hold information, and creating the conditions by which they are motivated to implement projects in a holistic manner [Rasul and Rogger, 2016; Rasul et al., 2018*a*; Bandiera et al., 2018]. Emphasizing measurement may improve officials' information but skew their incentives so to direct efforts towards those tasks that can best be measured. Broadly, further evidence is required on the tradeoffs of different management practices, and more broadly incentive environments, key to public service effectiveness.

[49]The fact that many private organizations are concerned with market outcomes is a key difference here. In the classical conception of the competitive market, information is created and shared organically as part of the market process. As sellers post prices for a good and consumers demand goods at those prices, the demand and supply curves, and thus equilibrium, of the corresponding market become public information. Private sector bureaucrats can access prices relevant to their work at close to zero marginal cost. Public bureaucracies do not have analogous information creation mechanisms, but rather individuals must undertake costly actions to acquire and absorb information, which then feed into allocation decisions.

Appendix

Theory Appendix: Deriving Implications of Aghion and Tirole (1997)

In the framework of Aghion and Tirole (1997), the principal hires an agent to acquire information regarding a project that the agent will implement. There are multiple projects, subscripted k, which each provide the principal with monetary gain B_k and the agent with b_k.[50] When no project is implemented $B_0 = b_0 = 0$. There is also one project that delivers a sufficiently negative payoff that an uninformed player will prefer inaction to recommending a project. The principal's preferred project yields monetary gain B and the agent's preferred project delivers b.

If the principal's preferred project is chosen, then the expected benefit to the agent is βb and if the agent's preferred project is chosen, the expected benefit to the principal is αB. α and β are (exogenous) 'congruence' parameters and belong to $(0, 1]$. The principal is risk-neutral and has utility $B_k - w$ if project k is chosen, where w is the wage paid to the agent; the agent is risk-averse and protected by limited liability ($w \geq 0$) and has utility $u(w) + b_k$, with $u' > 0$, $u'' < 0$. Projects are incomplete, since projects cannot be described and contracted on ex ante.

To focus on the information-acquisition trade-offs, it is assumed that the agent is infinitely risk-averse so that she does not respond to monetary incentives. The information game arises because the nature of the payoffs of the projects is initially unknown to both the principal and the agent. Both agents can exert costly effort to acquire the information about the payoffs of the projects. By exerting effort (e) at a cost of $g_A(e)$, the agent perfectly learns the payoffs of all candidate projects with probability e and learns nothing with the complement. By exerting effort E at a cost of $g_P(E)$, the principal becomes perfectly informed about the payoffs of all candidate projects with probability E and learns nothing with probability $1 - E$. In this paper we focus on the simultaneous variant of this game where the acquisition of information is made contemporaneously with the agents and $g_i' \geq 0$, $g_i'' > 0$, $g_i(0) = 0$, $g_i'(0) = 0$, $g_i'(1) = \infty$, $i = \{A, P\}$.

AT present cases of the game under: (i) P-Authority (the principal can always overrule the agent); and (ii) A-Authority (the principal cannot overrule the agent). The timing is: (i) The principal proposes a contract that allocates formal authority either to the principal or to the agent over the future choice of projects; (ii) the parties privately gather information about all projects' payoffs with costly effort; (iii) the party without formal authority communicates to the formal-authority party a subset of the

[50] For ease of comparison to AT, we stick to their notation in this appendix. For clarity with the notation used in the main part of the paper, U_{ii} is B (b) for the principal (agent), Uij is αB (βb) for the principal (agent), and E (e) is the probability of the principal (agent) being informed (perfectly learning the payoffs of all candidate projects).

relevant projects' payoffs he has learned; (iv) the formal-authority party picks a project (or none) on the basis of their information and the information communicated by the other.

Aghion and Tirole (1997): P-Authority

This sub-section presents the decision problem and equilibrium outcome of Aghion and Tirole (1997) under P-Authority – i.e. in the case where the principal has the formal authority and can always overrule the agent. The section below presents the case under A-Authority: the case where the agent has formal authority and she cannot be overruled by the principal.

In this case, the principal can always overrule the agent, and always does so if she is informed and if the agent recommends a less-than-optimum project for her. If she is uninformed and the agent is informed, she optimally rubber-stamps the proposal of the agent, since the congruence parameters are always positive. Rubber-stamping is akin to the agent having 'real authority'. The utility functions of the principal, u_P, and the agent, u_A, are as follows:

$$\begin{aligned} u_P &= EB + (1-E)e\alpha B - g_P(E) \\ u_A &= E\beta B + (1-E)eb - g_A(e) \end{aligned}$$

The first-order conditions can then be written:

$$\begin{aligned} B - \alpha B g_A'^{-1}(b - Eb) - g_P'(E) &= 0 \\ b - b g_P'^{-1}(B - e\alpha B) - g_A'(e) &= 0 \end{aligned} \tag{18}$$

Then by defining $G(E;\alpha,B,b) = B - \alpha B g_A'^{-1}(b-Eb) - g_P'(E)$ and $H(e;\alpha,B,b) = b - b g_P'^{-1}(B - e\alpha B) - g_A'(e)$, with G and H continuous in all arguments, and assuming that there exists some $e, E \in [0,1]$ such that $\frac{\partial G}{\partial E} \neq 0$ and $\frac{\partial H}{\partial e} \neq 0$, we can use the implicit function theorem to understand the comparative statics around reducing the marginal cost (g_P' and g_A') of information acquisition:

$$\begin{aligned} \frac{\partial E}{g_P'} &= \frac{-\partial G / \partial g_P'}{-\partial G / \partial E} \\ &= \frac{1}{\frac{\alpha b B}{g_A''} - g_P''} \end{aligned} \tag{19}$$

$$\frac{\partial e}{g'_A} = \frac{-\partial H/\partial g'_A}{-\partial H/\partial e} = \frac{1}{\frac{\alpha bB}{g''_P} - g''_A} \tag{20}$$

The first-order conditions are local maximizers of the utility function and provide stable equilibria if the second-order condition holds:

$$\alpha bB - g''_P g''_A < 0$$

This gives the intuitive result that both the optimal effort of the principal and of the agent to acquire information increases as the marginal cost decreases. Similarly, the effect of a change in the principal's marginal cost of effort on the optimal effort allocation of the agent and vice versa can be derived:

$$\frac{\partial E}{g'_A} = \frac{\alpha B}{g''_A g''_P - \alpha bB} \tag{21}$$

With the arguments of $g''_A = g''_A(\frac{B - g'_P(E)}{\alpha B})$ and $g''_P = g''_P(E)$.

$$\frac{\partial e}{g'_P} = \frac{b}{g''_A g''_P - \alpha bB} \tag{22}$$

With the argument of $g''_P = g''_P(\frac{b - g'_A(e)}{b})$ and $g''_A = g''_A(e)$.

Then, for simplicity, ignoring the arguments of g''_P and g''_A, and therefore assuming that these functions change sufficiently slowly within their arguments to alter the comparative statics, the total differentiation of a simultaneous change in the marginal cost of effort of the principal and the marginal cost of effort of the employee can be written:

$$TD_P = \Delta g'_P \Big[\frac{g''_A - b}{\alpha bB - g''_P g''_A}\Big] + \Delta g'_A \Big[\frac{g''_P - \alpha B}{\alpha bB - g''_P g''_A}\Big] \tag{23}$$

From Moulin [1984], we know that in a strictly concave game a global equilibrium is guaranteed if:[51]

$$\Big|\frac{\partial u_i^2}{\partial x_i^2(x)}\Big| > \Sigma_{\substack{j=1,\ldots,N \\ j \neq i}} \Big|\frac{\partial u_i}{\partial x_i \partial x_j}(x)\Big|$$

In this game, this requires:

[51] A concave game is one in which the payoff functions of the players are strictly concave in the players' own actions.

$$g''_A - b \geq 0$$

and

$$g''_P - \alpha B \geq 0$$

If we impose these two conditions, then the second-order conditions naturally hold as the numerators in equation 23 are both weakly positive and the denominator is strictly negative.

Taking into account the full range of effects implies that the total effect of a change in the marginal costs of effort of both the principal and agent will be ambiguously higher or lower as the parameters change. This is a natural result of the strategic substitutability between the agent's effort and the principal's effort. Imposing two additional assumptions, (i) that the third-order differentials of the cost functions g'''_P and g'''_A are sufficiently small so that marginal changes in their arguments can be ignored, and (ii) that the direct effects of the treatment override indirect effects associated with free-riding concerns, leads to the following predictions:[52]

- A reduction in the marginal cost of effort of the principal and the employee will increase the total effort exerted to acquire information[53]
- As the congruence α increases, the effect of changing the marginal cost of effort on subsequent effort will be smaller (a smaller treatment effect would be expected) – this occurs through $\Delta g'_A$
- As the benefit to the principal B increases, the effect of changing the marginal cost of effort on subsequent effort will be smaller (a smaller treatment effect would be expected) – this occurs through $\Delta g'_A$
- As the benefit to the agent b increases, the effect of changing the marginal cost of effort on subsequent effort will be smaller (a smaller treatment effect would be expected) – this occurs through $\Delta g'_P$

[52]The 'direct' effect is the response of the agent (principal) to a reduction in the marginal cost of her effort, while the 'indirect' effect or free-riding occurs through the response of the agent (principal) to a reduction in the marginal cost of the principal's (agent's) effort. These can be observed by differentiating TD_P with respect to the parameters of interest (α, b, B).

[53]Note that this is guaranteed under the stability conditions assumed in AT. The additional assumption regarding the overriding of free-riding concerns is not necessary for this prediction under this framework but allows generalizability to cases where free-riding concerns exist and the stability conditions may not hold.

Aghion and Tirole (1997): A-Authority

This sub-section presents the case under A-Authority: the case where the agent has formal authority – she cannot be overruled by the principal.

In this case, if the agent is informed, she will pick her favored project and cannot be overruled by the principal. If she is uninformed and the principal is informed, she accepts the recommendation of the principal.

$$u_P^d = e\alpha B + (1-e)EB - g_P(E)$$
$$u_a^d = eb + (1-e)E\beta b - g_A(e)$$

The first-order conditions can be written:

$$\begin{aligned} B - g_A'^{-1}(b(1-E\beta))B - g_P'(E) = 0 \\ b - g_P'^{-1}((1-e)B)\beta b - g_A'(e) = 0 \end{aligned} \tag{24}$$

Following the same rationale as above, using the implicit function theorem, the comparative statics of changes in the marginal cost of information on the equilibrium effort allocations can be derived as follows:

$$\begin{aligned} \frac{\partial E}{\partial g_P'} &= \frac{g_A''}{bB\beta - g_P''g_A''} \\ \frac{\partial e}{\partial g_A'} &= \frac{g_P''}{bB\beta - g_P''g_A''} \\ \frac{\partial E}{\partial g_A'} &= \frac{-B}{bB\beta - g_P''g_A''} \\ \frac{\partial e}{\partial g_P'} &= \frac{-b\beta}{bB\beta - g_P''g_A''} \end{aligned}$$

The total effect of a marginal change in the marginal cost of effort for both the principal and the agent simultaneously follows from the total differentiation:

$$TD_A = \Delta g_P'\Big[\frac{g_A'' - \beta b}{bB\beta - g_A''g_P''}\Big] + \Delta g_A'\Big[\frac{g_P'' - B}{bB\beta - g_A''g_P''}\Big] \tag{25}$$

Sufficient conditions for global stability of the equilibrium, using the conditions from Moulin (1984), require $g_P''(E) - B \geq 0$ and $g_A''(e) - b\beta \geq 0$. Imposing these conditions and the second-order condition $bB\beta - g_A''g_P'' < 0$; and then imposing that (i) the third-order differentials of the cost functions g_P''' and g_A''' are sufficiently small so

that marginal changes in their arguments can be ignored, and (ii) that the direct effects of the treatment override indirect effects associated with free-riding concerns, leads to the following predictions:[54]

- A reduction in the marginal cost of effort of the principal and the employee will increase the total effort exerted to acquire information
- As the congruence β increases, the effect of changing the marginal cost of effort on subsequent effort will be smaller (a smaller treatment effect would be expected) – this functions through $\Delta g'_P$
- As the benefit to the agent b increases, the effect of changing the marginal cost of effort on subsequent effort will be smaller (a smaller treatment effect would be expected) – this functions through $\Delta g'_P$
- As the benefit to the principal B increases, the effect of changing the marginal cost of effort on subsequent effort will be smaller (a smaller treatment effect would be expected) – this functions through $\Delta g'_A$

Data Appendix: Creating Measures of Bureaucratic Error

This section describes in detail the creation of the error variables. For each individual surveyed, we obtained their subjective assessments ('claims') of demographic characteristics of their constituency, of the sector-specific variables relating to their constituency, and of the variables relating to the public service rules.

Where individuals stated that they did not know the answer at all, a value of 0 was entered to signal that the respondent did not know the answer to a particular question at all. Of the 891 individuals who make up our common-claims sample used in the core analysis, 275 (31%) of individuals respond 'Don't Know' to at least one of the questions asked. 181 individuals (20%) respond 'Don't Know' to all of the questions in the information module. There is, therefore, selection into the sample of respondents who provide non-zero estimates. Regressing individual characteristics on a binary variable that takes the value 1 if a respondent answers 'Don't Know' to all questions, we find evidence that district officials are less likely to do so. We make two comments regarding the selection into the error sample: firstly, it seems a reasonable assumption that the probability of responding 'Don't Know' is positively associated with the most extreme (latent) errors, and hence our estimates of the difference across

[54]The 'direct' effect is the response of the agent (principal) to a reduction in the marginal cost of her effort, while the 'indirect' effect or free-riding occurs through the response of the agent (principal) to a reduction in the marginal cost of the principal's (agent's) effort. These can be observed by differentiating TD_P with respect to the parameters of interest (α, b, B).

tiers can be interpreted as a lower bound. Secondly, the error data provide consistent estimates of the accuracy of information that is present within a public office in Ethiopia – for example, if one entered a public office at random in Ethiopia and asked about a particular service delivery indicator, the information that one would receive would be reflected within the error data we use for our regressions. To ensure that these 0 claims do not affect the analysis on the errors, a missing value is inserted in place of the 0 for the error variables. Furthermore, a 'Don't Know' binary indicator is created that takes the value 1 whenever a respondent does not know the answer to a question.

For each statistic we study, we subtract the corresponding benchmark data sourced from survey or administrative data from the claim of the respondent. For questions related to the rules of the service, which are uniform across all respondents regardless of sector and tier, the claim is subtracted from the benchmark data and the absolute value generated. For the demographic and sector claims, the format of these differ slightly by tier. District officials were asked directly about their district, and the relevant benchmark indicators subtracted from these responses. The regional officials were asked about the 25th, 50th and 75th percentile districts of individual variables separately. The federal officials were asked, "Which region, out of the following options, do you work on most/think that you know best/feel most comfortable answering questions about?" They were then asked about the 25th, 50th and 75th percentile districts of individual variables within that region. This ensures that each official is responding to questions on the districts that they are most familiar with, and work most closely on. Federal officials across sectors chose the full distribution of regions, and the distribution of choices closely matches population and district numbers.

The raw errors vary significantly, as documented in the main text. The distribution of the errors suggests that civil servants, on average, tend to overestimate the value of the indicator, with the average error of all but three of the indicators we study being positive (the three indicators with a negative actual error are: the proportion of rural inhabitants; the primary pupil-class ratio; and the hectares of land used for pastoral purposes). Since it is difficult to compare errors across different indicators, the overall skewness coefficient for the z-score of the actual error is 0.4. A test of skewness relative to the normal distribution is rejected at the usual levels. Individually, the indicators tend to be positively skewed (have distributions with longer right tails). Six of the 20 indicators have a negative skewness coefficent: proportion of rural inhabitants (-3.02); the antenatal care rate (-0.49); the rate of births delivered by a skilled attendant (-0.22); the share of households living on subsistence agriculture (-0.73); the proportion of income from regional block grants (0.64); and the proportion of income from own sources (-0.56).

In the core analysis, we aim to compare the claims of officials at multiple tiers of

government relating to the same district. We therefore focus on districts for which we have subjective assessments from each tier of government. This requires a district to be one of the 25th/50th/75th percentiles for at least one of the indicators we study, since these are the districts over which federal and regional officials made assessments. We therefore drop roughly a third (24) of the 66 districts on which we collected data to ensure that each claim has counterparts at all three tiers of government.

We create a data set where the unit of observation is a claim of a bureaucrat about a district. As described, each federal and regional official is asked about three districts for each indicator (the 25th/50th/75th percentiles of that indicator), and so each bureaucrat-indicator-percentile is an observation for these officials. The claim of the district official about a specific indicator in her district is a single observation. Each response is associated with the corresponding district census ID.

In a very small number of cases, there seems to have been a data entry issue, where additional digits were included or excluded in the recording of responses to non-fractional variables. This leads us to some errors that are off by a very significant factor (of 94 at the 95th percentile of errors). These are not reflective of the wider variation in the data. Two-thirds of absolute errors are less than 1 in proportional terms. The most extreme values seem outside the realm of feasiblity for even the least informed bureaucrat. Since a plausible explanation is that these outliers reflect data entry issues rather than outsized errors, we therefore censor the top 5% of raw errors for each of the variables. The interpretation of the paper is therefore analysis of the lowest 95% of errors that public officials make, rather than of the top 5%.

Absolute values are then taken of all errors. At this stage, we have a data set that is not comparable across, but only within indicators. Making an error of 0.5 in the percentage of rural inhabitants is large for that indicator, but tiny for the population indicator. We therefore undertake a series of transformations of the absolute errors to make them comparable across indicators. We take three approaches.

Our core measure is created by constructing z-scores that are the value of the absolute error minus the mean value across the distribution of absolute errors within the indicator, all divided by the standard deviation of absolute errors across the distribution of absolute errors within the indicator. The interpretation of this z-score is a unitless measure of how many standard deviations away from the benchmark data an official's assessment is. Comparisons across indicators therefore take into account the empirical distribution of errors for each indicator. They compare the scale of an official's error relative to the errors made by all other officials in our data set.

One issue with the above measure is that it does not take into account potential skewness of the underlying indicators. Table A2 provides information on the Pearson moment coefficient of skewness for each of the indicators. It is evident that the non-

fraction indicators exhibit right-skew, with positive coefficients larger than 1 in all cases. For non-fraction indicators, a test for skewness relative to the normal distribution is conducted based on D'agostino et al. [1990] and Royston [1992]. The test statistic is approximately normal under the null hypothesis and allows for standard hypothesis tests. Thus, if the p-value associated with the test statistic is less than or equal to 0.05, we classified the variable as significantly skewed and replace the indicator with its log, consistent with Duggan and Scott Morton [2006]. For those that are not significantly skewed, no transformation is applied. The 'skewness-adjusted' z-score is calculated on this set of values.

A third approach is to use the ordinal rank of the absolute error within the entire distribution of absolute errors within the indicator (again consistent with Duggan and Scott Morton [2006]. We rank all errors according to size, give each a corresponding rank, and divide these ranks by the maximum rank within the distribution of absolute errors within the indicator. The rank-proportion scores are therefore free of distributional assumptions. They can be interpreted as the percentile of error made by an official relative to other officials in our data set.

Table 23: Characteristics of Civil Servants

Table 1: Characteristics of Civil Servants

Means and standard deviations

	(1) All	(2) Federal	(3) Regional	(4) Local
Organizational Characteristics				
Number of organizations	382	5	54	323
Number of bureaucrats per organization	5.71	56.60	8.67	4.43
Number of managers per organization	1.22	6.80	2.57	0.90
Number of employees per organization	3.58	49.20	5.41	2.56
Span of control (employees per manager)	7.79	27.28	15.98	5.97
Official's Characteristics				
Number of bureaucrats	1831	280	431	1120
Age	35.30	35.87	38.10	34.07
	[8.85]	[9.10]	[8.74]	[8.58]
Years in position	2.70	2.29	2.98	2.69
	[2.74]	[2.88]	[2.87]	[2.64]
Years in organization	7.34	5.80	7.97	7.48
	[7.13]	[6.47]	[7.27]	[7.18]
Years in civil service	13.13	12.87	15.49	12.29
	[8.88]	[8.98]	[9.00]	[8.65]
Number of service organizations worked in	2.78	3.02	3.08	2.55
	[2.01]	[2.07]	[2.40]	[1.73]
Grade	7.93	7.53	10.89	6.90
	[13.06]	[9.92]	[17.97]	[11.21]
Education [undergraduate degree=1]	0.82	0.96	0.90	0.76
	[0.38]	[0.19]	[0.30]	[0.43]
Education [masters degree=1]	0.11	0.35	0.22	0.01
	[0.32]	[0.48]	[0.42]	[0.12]
Gender [female=1]	0.20	0.27	0.23	0.18
	[0.40]	[0.44]	[0.42]	[0.38]
Manager	0.25	0.12	0.32	0.26
	[0.44]	[0.33]	[0.47]	[0.44]

Notes: Standard deviations are in parentheses. The unit of observation for the first panel is the civil servant and for the second panel the organization. Only employees and managers responded to the information module and hence only these individuals are used for the statistics in the first panel. These statistics are based on the survey sample and not on administrative data. The span of control is taken from the survey question "How many personnel do you manage?"; the average is taken for an organization if there is more than one manager; and the organization-level statistics are reported (each organization has an equal weight in the `By Organization' statistics). `Manager' refers to managers within the organization. Grade is the official civil service grade of the professional civil servant, ranging from 1 to 17 in the sample. The 2013-14 National Civil Service Human Resource Statistics Abstract (Ministry of Civil Service) records the overall female employee percentage to be 35%; this statistic includes frontline staff. Figures are rounded to two decimal places.

Table 2: Hierarchy and Information

Dependent Variable: A binary indicator of whether respondent doesn't know in columns (1) and (2); Z-score of errors across all policy domains in columns (3) to (10)
OLS Estimates
Standard Errors: Clustered at the district government-sector level

	(1) Don't know Unconditional	(2) Don't know Conditional	(3) Unconditional Decentralized	(4) Unconditional Manager	(5) Unconditional Hierarchies	(6) Sector FEs	(7) District government FEs	(8) Indicator FEs	(9) Percentile FEs	(10) Individual Characteristics
District Government [Yes=1]	-0.17***	-0.17***	-0.27***		-0.27***	-0.28***	-0.30***	-0.31***	-0.31***	-0.33***
	[0.034]	[0.033]	[0.084]		[0.084]	[0.084]	[0.086]	[0.087]	[0.087]	[0.087]
Manager [Yes=1]	0.011	0.0040		0.011	0.023	0.020	0.014	0.016	0.015	0.034
	[0.031]	[0.029]		[0.066]	[0.066]	[0.067]	[0.065]	[0.066]	[0.066]	[0.069]
Sector fixed effects	No	Yes	No	No	No	Yes	Yes	Yes	Yes	Yes
District government fixed effects	No	Yes	No	No	No	No	Yes	Yes	Yes	Yes
Indicator fixed effects	No	Yes	No	No	No	No	No	Yes	Yes	Yes
Percentile fixed effects	No	Yes	No	No	No	No	No	No	Yes	Yes
Individual controls	No	Yes	No	No	No	No	No	No	No	Yes
Adjusted R-squared	0.022	0.13	0.012	-0.00066	0.012	0.0099	0.059	0.067	0.072	0.069
Observations [clusters]	2144 [98]	2144 [98]	1455 [92]	1455 [92]	1455 [92]	1455 [92]	1455 [92]	1455 [92]	1455 [92]	1455 [92]
Central government respondents [percenta	594 [67%]	594 [67%]	429 [62%]	429 [62%]	429 [62%]	429 [62%]	429 [62%]	429 [62%]	429 [62%]	429 [62%]
District government respondents [percenta	297 [33%]	297 [33%]	258 [38%]	258 [38%]	258 [38%]	258 [38%]	258 [38%]	258 [38%]	258 [38%]	258 [38%]

Notes: *** denotes significance at 1%, ** at 5%, and * at 10% level. Clustered standard errors are in parentheses. Standard errors clustered at the district government-sector interacted level. All columns report OLS estimates. The unit of analysis in all columns is the claim. The sample used for the analysis is the set of respondents that make claims about the same districts. The dependent variable is a binary indicator for whether the respondent could not provide a claim in columns (1) and (2); the z-score of absolute error across all policy domain items in columns (3) to (12). Sector fixed effects are indicators for whether the claim is within an agriculture, education, health, revenue, or trade organization. Percentile fixed effects are indicators for the 75th, 50th or 25th percentile for the respective indicator for the respective region. District government fixed effects identify the district that the respondent is making claims about. Indicator fixed effects control for the service delivery indicator within which the claim is made. Manager is equal to one if the claim is from a director. Individual characteristics included in the specification: tenure in the civil service of the respondent making the claim; tenure in the position; tenure in the organization; an indicator for whether the respondent is female; an indicator for whether the respondent holds an undergraduate degree; an indicator for whether the respondent holds a masters degree. Figures are rounded to two significant figures.

Table 3: Information, Incentives and Alignment

Dependent Variable: Z-score of errors across all policy domains in all columns

OLS Estimates

Standard Errors: Clustered at the district government-sector level

	Incentives				
	(1) Management practices	(2) Noise controls	(3) Combined	(4) Interaction Decentralized	(5) Interaction Manager
District Government [Yes=1]			-0.27*	-0.34**	-0.27*
			[0.16]	[0.16]	[0.16]
Manager [Yes=1]	0.011	-0.026	-0.029	-0.021	-0.035
	[0.070]	[0.076]	[0.076]	[0.077]	[0.076]
Management Practice: Monitoring	-0.17**	-0.21***	-0.20***	-0.14*	-0.17**
	[0.067]	[0.067]	[0.067]	[0.080]	[0.067]
Management Practice: Other	0.28***	0.31***	0.32***	0.32***	0.31***
	[0.097]	[0.11]	[0.11]	[0.10]	[0.11]
District Government x Management Practice: Monitoring				-0.19	
				[0.14]	
Manager x Management Practice: Monitoring					-0.10
					[0.090]
Sector, indicator, percentile fixed effects	Yes	Yes	Yes	Yes	Yes
District government fixed effects	No	No	No	No	No
Region fixed effects	Yes	Yes	Yes	Yes	Yes
Individual controls	Yes	Yes	Yes	Yes	Yes
Marginal costs and noise controls	No	Yes	Yes	Yes	Yes
Adjusted R-squared	0.034	0.061	0.063	0.064	0.063
Observations [clusters]	1455 [92]	1455 [92]	1455 [92]	1455 [92]	1455 [92]
Central government respondents [percentage]	429 [62%]	429 [62%]	429 [62%]	429 [62%]	429 [62%]
District government respondents [percentage]	258 [38%]	258 [38%]	258 [38%]	258 [38%]	258 [38%]

Notes: *** denotes significance at 1%, ** at 5%, and * at 10% level.Clustered standard errors are in parentheses.Standard errors clustered at the district government-sector interacted level, allowing for error correlation within regions, sectors, and districts.All columns report OLS estimates.The unit of analysis in all columns is the claim.The tiers of government are grouped into District Government and Centralized Government which incorporates both Federal and Regional tiers, hence the omitted category for tier in the regressions is Centralized Government.The sample used for the analysis is the set of respondents that make claims about the same districts.The dependent variable is the z-score of absolute errors across all policy domains.Sector fixed effects are indicators for whether the claim is within an agriculture, education, health, revenue, or trade organization.Percentile fixed effects are indicators for the 75th, 50th or 25th percentile for the respective indicator for the respective region.District government fixed effects identify the district that the respondent is making claims about.Indicator fixed effects control for the service delivery indicator within which the claim is made.Management practice: Monitoring is the organization-average z-score for the monitoring practice of management as per Bloom and Van Reenen (2007); Management practice: Other is the mean of the organization average z-scores for targeting, incentives, roles, flexibility, staffing, and staff involvement. Manager is equal to one if the claim is from a manager.Individual characteristics included in the specification:tenure in the civil service of the respondent making the claim;tenure in the position; tenure in the organization;an indicator for whether the respondent is female; an indicator for whether the respondent holds an undergraduate degree; an indicator for whether the respondent holds a masters degree; and an indicator for whether the respondent is a manager. Marginal costs include the proportion of managers in the organization that state that a management information system is in place; the organization average response to "In what proportion of projects would you say information flows effectively between individuals in the same organization?"; the organization average response to "In what proportion of projects would you say Information flows effectively between organizations?"; the organization average response of managers to "How many personnel do you manage?"; the organization average response of employees to "How many people would you say regularly give you tasks as part of your formal work duties?";alignment (the z-score of the row mean of the organization-average response to "To what extent do you believe that your organization's mission is aligned to your own mission?" and the organization proportion of selecting "I am contributing to fulfilling that mission on an everyday basis" to "How consistent with your mission are the various tasks and activities assigned to you on a day-to-day basis?").Noise controls include the time of day of the survey, day-of-survey fixed effects, enumerator fixed effects, an indicator for the enumerator's subjective assessment of the quality of the interview; an indicator for the duration of the interview (decile of duration fixed effects); controls for the average tenure of the managers in the organization and the proportion of female managers in the organization.Figures are rounded to two significant figures.

Table 4: Information Experiment

Dependent Variable: Z-score of errors across policy domains in columns (1) to (4) and (7); rank of error (proportion) in column (5); and skewness-adjusted z-score in column (6)

OLS estimates in all columns

Standard Errors: Clustered at the district government-sector level in columns (1) to (6) and at the region-sector level in column (7)

	(1) Treatment only	(2) Conditional Treatment Effect	(3) Baseline with treatment	(4) Management practices	(5) Rank Error (Proportion)	(6) Skewness-Adjusted Z-Score Error	(7) Clustering at Region-Sector Level
Treated District [Yes=1]	0.053	-0.15*	-0.28***	-0.23**	-0.074**	-0.20*	-0.23**
	[0.087]	[0.085]	[0.098]	[0.11]	[0.032]	[0.11]	[0.11]
P-value of coefficient clustered at district government-se	*{0.54}*	*{0.078}*	*{0.0048}*	*{0.036}*	*{0.022}*	*{0.086}*	*{0.048}*
P-value of coefficient clustered at district level	*(0.51)*	*(0.057)*	*(0.0031)*	*(0.033)*	*(0.024)*	*(0.053)*	*(0.033)*
District Government [Yes=1]			-0.44**	-0.41**	-0.11**	-0.39**	-0.41**
			[0.17]	[0.17]	[0.048]	[0.17]	[0.16]
Manager [Yes=1]			-0.021	-0.026	-0.011	-0.0074	-0.026
			[0.077]	[0.076]	[0.020]	[0.066]	[0.080]
Management Practice: Monitoring				-0.18**	-0.018	-0.027	-0.18***
				[0.069]	[0.024]	[0.088]	[0.064]
Management Practice: Other				0.28**	0.038	0.090	0.28*
				[0.11]	[0.038]	[0.13]	[0.14]
Sector, region, percentile, indicator fixed effects	No				Yes		
Individual controls	No				Yes		
Marginal costs and noise controls	No				Yes		
Adjusted R-squared	-0.00018	0.057	0.062	0.066	0.092	0.10	0.066
Number of observations [clusters]	1455 [92]	1455 [92]	1455 [92]	1455 [92]	1455 [92]	1455 [92]	1455 [38]

Notes: *** denotes significance at 1%, ** at 5%, and * at 10% level. All columns report OLS regression coefficients and standard errors in brckets. Standard errors in brackets. Standard errors clustered at the district government-sector interacted level in columns (1) to (6) and at the region-sector interacted level in column (7). P-values are reported in curly brackets. P-values associated with clustering at the district level are reported in round brackets. The unit of analysis in all columns is the claim. The sample used for the analysis is the set of respondents that make claims about the same districts. The dependent variable is the z-score error for any claim made within any policy domain in columns (1) to (4) and column (7). The dependent variable is the rank of the absolute error in terms of the proportion of the total number of claims for the indicator in column (5). The dependent variable is the skewness-adjusted z-score in column (6). The tiers of government are grouped into District Government and Centralized Government which incorporates both Federal and Regional tiers, hence the omitted category for tier in the regressions is Centralized Government. Treated district is an indicator equal to one if the district received an information package. Sector fixed effects are indicators for whether the claim is within an agriculture, education, health, revenue, or trade organization. Percentile fixed effects are indicators for the 75th, 50th or 25th percentile for the respective indicator for the respective region. Indicator fixed effects are indicator variables for which specific variable the claim is being made over. Management practice: Monitoring is the organization-average z-score for the monitoring practice of management as per Bloom and Van Reenen (2007); Management practice: Other is the mean of the organization average z-scores for targeting, incentives, roles, flexibility, staffing, and staff involvement. Manager is equal to one if the claim is from a manager. Individual characteristics included in the specification: tenure in the civil service of the respondent making the claim; tenure in the position; tenure in the organization; an indicator for whether the respondent is female; an indicator for whether the respondent holds an undergraduate degree; an indicator for whether the respondent holds a masters degree. Marginal costs include the proportion of managers in the organization that state that a management information system is in place; the organization average response to "In what proportion of projects would you say information flows effectively between individuals in the same organization?"; the organization average response to "In what proportion of projects would you say Information flows effectively between organizations?"; the organization average response of managers to "How many personnel do you manage?"; the organization average response of employees to "How many people would you say regularly give you tasks as part of your formal work duties?"; alignment (the z-score of the row mean of the organization-average response to "To what extent do you believe that your organization's mission is aligned to your own mission?" and the organization proportion of selecting "I am contributing to fulfilling that mission on an everyday basis" to "How consistent with your mission are the various tasks and activities assigned to you on a day-to-day basis?"). Noise controls include the time of day of the survey, day-of-survey fixed effects, enumerator fixed effects, an indicator for the enumerator's subjective assessment of the quality of the interview; an indicator for the duration of the interview (decile of duration fixed effects); controls for the average tenure of the managers in the organization and the proportion of female managers in the organization. Figures are rounded to two decimal places or two significant figures.

Table 5: Heterogeneous Treatment Effects

Dependent Variable: Z-score of errors across policy domains
OLS estimates in all columns
Standard Errors: Clustered at the district government-sector level

	(0) Baseline treatment effect	(1) Interaction: district government	(2) Interaction: manager	(3) Interaction: monitoring	(4) Interaction: other management practices
Treated District [Yes=1]	-0.23**	-0.025	-0.20*	-0.26**	-0.34***
	[0.11]	[0.19]	[0.12]	[0.10]	[0.11]
District Government [Yes=1]	-0.41**	-0.089	-0.41**	-0.46***	-0.43***
	[0.17]	[0.28]	[0.17]	[0.16]	[0.16]
Manager	-0.026	-0.032	0.068	-0.018	-0.015
	[0.076]	[0.075]	[0.13]	[0.076]	[0.076]
Management Practice: Monitoring	-0.18**	-0.18**	-0.18**	-0.43***	-0.18***
	[0.069]	[0.073]	[0.069]	[0.11]	[0.065]
Management Practice: Other	0.28**	0.30**	0.28**	0.30***	0.023
	[0.11]	[0.12]	[0.11]	[0.11]	[0.14]
Treated District x District Government [Yes=1]		-0.37			
		[0.25]			
Treated District x Manager			-0.12		
			[0.13]		
Treated District x Management Practice: Monitoring				0.34***	
				[0.13]	
Treated District x Management Practice: Other					0.38***
					[0.14]
Sector, region, percentile, indicator fixed effects			Yes		
Individual controls			Yes		
Marginal costs and noise controls			Yes		
Adjusted R-squared	0.066	0.067	0.065	0.071	0.071
Number of observations [clusters]	1455 [92]	1455 [92]	1455 [92]	1455 [92]	1455 [92]

Notes: *** denotes significance at 1%, ** at 5%, and * at 10% level. All columns report OLS regression coefficients and standard errors in brckets. Standard errors clustered at the district government-sector interacted level. The unit of analysis in all columns is the claim. The sample used for the analysis is the set of respondents that make claims about the same districts. The dependent variable is the z-score error for any claim made within any policy domain. The tiers of government are grouped into District Government and Centralized Government which incorporates both Federal and Regional tiers, hence the omitted category for tier in the regressions is Centralized Government. Treated district is an indicator equal to one if the district received an information package. Sector fixed effects are indicators for whether the claim is within an agriculture, education, health, revenue, or trade organization. Percentile fixed effects are indicators for the 75th, 50th or 25th percentile for the respective indicator for the respective region. Indicator fixed effects are indicator variables for which specific variable the claim is being made over. Management practice: Monitoring is the organization-average z-score for the monitoring practice of management as per Bloom and Van Reenen (2007); Management practice: Other is the mean of the organization average z-scores for targeting, incentives, roles, flexibility, staffing, and staff involvement. Manager is equal to one if the claim is from a manager. Individual characteristics included in the specification: tenure in the civil service of the respondent making the claim; tenure in the position; tenure in the organization; an indicator for whether the respondent is female; an indicator for whether the respondent holds an undergraduate degree; an indicator for whether the respondent holds a masters degree; an indicator if the respondent is a manager. Marginal costs include the proportion of managers in the organization that state that a management information system is in place; the organization average response to "In what proportion of projects would you say information flows effectively between individuals in the same organization?"; the organization average response to "In what proportion of projects would you say Information flows effectively between organizations?"; the organization average response of managers to "How many personnel do you manage?"; the organization average response of employees to "How many people would you say regularly give you tasks as part of your formal work duties?"; alignment (the z-score of the row mean of the organization-average response to "To what extent do you believe that your organization's mission is aligned to your own mission?" and the organization proportion of selecting "I am contributing to fulfilling that mission on an everyday basis" to "How consistent with your mission are the various tasks and activities assigned to you on a day-to-day basis?"). Noise controls include the time of day of the survey, day-of-survey fixed effects, enumerator fixed effects, an indicator for the enumerator's subjective assessment of the quality of the interview; an indicator for the duration of the interview (decile of duration fixed effects); controls for the average tenure of the managers in the organization and the proportion of female managers in the organization. Figures are rounded to two decimal places or two significant figures.

Table 28: Surveyed Organisations

TA1: List of Surveyed Organizations

Tier of Governance	Region	Organization Name
Federal	-	Federal Ministries of Agriculture; Education; Health; Revenue; and, Trade
Regional	Addis Ababa	Addis Ababa City Administration Bureaus of Education; Health; Revenue; and, Trade
Regional	Afar	Afar Regional Bureaus of Agriculture; Education; Health; Revenue; and, Trade
Regional	Amhara	Amhara Regional Bureaus of Agriculture; Education; Health; Revenue; and, Trade
Regional	Benishangul Gumuz	Benishangul Gumuz Regional Bureaus of Agriculture; Education; Health; Revenue; and, Trade
Regional	Dire Dawa	Dire Dawa City Administration Bureaus of Agriculture; Education; Health; Revenue; and, Trade
Regional	Gambella	Gambella Regional Bureaus of Agriculture; Education; Health; Revenue; and, Trade
Regional	Harar	Harar Regional Bureaus of Agriculture; Education; Health; Revenue; and, Trade
Regional	Oromia	Oromia Regional Bureaus of Agriculture; Education; Health; Revenue; and, Trade
Regional	SNNPR	SNNPR Regional Bureaus of Agriculture; Education; Health; Revenue; and, Trade
Regional	Somali	Somali Regional Bureaus of Agriculture; Education; Health; Revenue; and, Trade
Regional	Tigray	Tigray Regional Bureaus of Agriculture; Education; Health; Revenue; and, Trade
District	Afar	Afar Awash Fentale Agriculture, Education, Health, Revenue, and Trade Offices
District	Afar	Afar Telalak Agriculture, Education, Health, Revenue, and Trade Offices
District	Afar	Afar Teru Agriculture, Education, Health, Revenue, and Trade Offices
District	Amhara	Amhara Awabel Agriculture, Education, Health, Revenue, and Trade Offices
District	Amhara	Amhara Basona Worana Agriculture, Education, Health, Revenue, and Trade Offices
District	Amhara	Amhara Borena (Former Debresina) Agriculture, Education, Health, Revenue, and Trade Offices
District	Amhara	Amhara Chefa Gula Agriculture, Education, Health, Revenue, and Trade Offices
District	Amhara	Amhara Dejen Agriculture, Education, Health, Revenue, and Trade Offices
District	Amhara	Amhara Enarj Enawaga Agriculture, Education, Health, Revenue, and Trade Offices
District	Amhara	Amhara Gidane Agriculture, Education, Health, Revenue, and Trade Offices
District	Amhara	Amhara Jabitahnan Agriculture, Education, Health, Revenue, and Trade Offices
District	Amhara	Amhara Jile Timuga Agriculture, Education, Health, Revenue, and Trade Offices
District	Amhara	Amhara Kutaber Agriculture, Education, Health, Revenue, and Trade Offices
District	Amhara	Amhara Simada Agriculture, Education, Health, Revenue, and Trade Offices
District	Benishangul Gumuz	Benishangul Gumuz Dibate Agriculture, Education, Health, Revenue, and Trade Offices
District	Benishangul Gumuz	Benishangul Gumuz Yasso Agriculture, Education, Health, Revenue, and Trade Offices
District	Gambella	Gambella Gambella Zuria Agriculture, Education, Health, Revenue, and Trade Offices
District	Gambella	Gambella Abobo Agriculture, Education, Health, Revenue, and Trade Offices
District	Oromia	Oromia Ale Agriculture, Education, Health, Revenue, and Trade Offices
District	Oromia	Oromia Amigna Agriculture, Education, Health, Revenue, and Trade Offices
District	Oromia	Oromia Arsi Negelle District government Agriculture, Education, Health, Revenue, and Trade Offices
District	Oromia	Oromia Babile Agriculture, Education, Health, Revenue, and Trade Offices
District	Oromia	Oromia Bako Tibe Agriculture, Education, Health, Revenue, and Trade Offices
District	Oromia	Oromia Begi Agriculture, Education, Health, Revenue, and Trade Offices
District	Oromia	Oromia Dedessa Agriculture, Education, Health, Revenue, and Trade Offices
District	Oromia	Oromia Digluna Tijo Agriculture, Education, Health, Revenue, and Trade Offices
District	Oromia	Oromia Gida Ayana (Gida Kiremu) Agriculture, Education, Health, Revenue, and Trade Offices
District	Oromia	Oromia Goro Gutu (Goro) Agriculture, Education, Health, Revenue, and Trade Offices
District	Oromia	Oromia Guduru Agriculture, Education, Health, Revenue, and Trade Offices
District	Oromia	Oromia Haro Maya Agriculture, Education, Health, Revenue, and Trade Offices
District	Oromia	Oromia Hitosa Agriculture, Education, Health, Revenue, and Trade Offices
District	Oromia	Oromia Jardega Jarte Agriculture, Education, Health, Revenue, and Trade Offices
District	Oromia	Oromia Jeldu Agriculture, Education, Health, Revenue, and Trade Offices
District	Oromia	Oromia Kofale Agriculture, Education, Health, Revenue, and Trade Offices
District	Oromia	Oromia Mesela Agriculture, Education, Health, Revenue, and Trade Offices
District	Oromia	Oromia Midaga Tola Agriculture, Education, Health, Revenue, and Trade Offices
District	Oromia	Oromia Nono Agriculture, Education, Health, Revenue, and Trade Offices
District	Oromia	Oromia Seru Agriculture, Education, Health, Revenue, and Trade Offices
District	Oromia	Oromia Siraro Agriculture, Education, Health, Revenue, and Trade Offices
District	Oromia	Oromia Tikur Enchini Agriculture, Education, Health, Revenue, and Trade Offices
District	Oromia	Oromia Wadera Agriculture, Education, Health, Revenue, and Trade Offices
District	Oromia	Oromia Were Jarso Agriculture, Education, Health, Revenue, and Trade Offices
District	SNNPR	SNNPR Amaro Agriculture, Education, Health, Revenue, and Trade Offices
District	SNNPR	SNNPR Analimo Agriculture, Education, Health, Revenue, and Trade Offices
District	SNNPR	SNNPR Basketo Agriculture, Education, Health, Revenue, and Trade Offices
District	SNNPR	SNNPR Benatsemay Agriculture, Education, Health, Revenue, and Trade Offices
District	SNNPR	SNNPR Bona Agriculture, Education, Health, Revenue, and Trade Offices
District	SNNPR	SNNPR Chere Agriculture, Education, Health, Revenue, and Trade Offices
District	SNNPR	SNNPR Dale Agriculture, Education, Health, Revenue, and Trade Offices
District	SNNPR	SNNPR Decha Agriculture, Education, Health, Revenue, and Trade Offices
District	SNNPR	SNNPR Doyo Gena Agriculture, Education, Health, Revenue, and Trade Offices
District	SNNPR	SNNPR Gomibora Agriculture, Education, Health, Revenue, and Trade Offices
District	SNNPR	SNNPR Hawassa Zuriya Agriculture, Education, Health, Revenue, and Trade Offices
District	SNNPR	SNNPR Kucha Agriculture, Education, Health, Revenue, and Trade Offices
District	SNNPR	SNNPR Shebedino Agriculture, Education, Health, Revenue, and Trade Offices
District	SNNPR	SNNPR Wenago Agriculture, Education, Health, Revenue, and Trade Offices
District	Somali	Somali Afdem Agriculture, Education, Health, Revenue, and Trade Offices
District	Somali	Somali Erer District government Agriculture, Education, Health, Revenue, and Trade Offices
District	Somali	Somali Harshin Agriculture, Education, Health, Revenue, and Trade Offices
District	Somali	Somali Jijiga Zuria District government Agriculture, Education, Health, Revenue, and Trade Offices
District	Somali	Somali Kebri Beyah Agriculture, Education, Health, Revenue, and Trade Offices
District	Tigray	Tigray Erob Agriculture, Education, Health, Revenue, and Trade Offices
District	Tigray	Tigray Gulo Mekeda Agriculture, Education, Health, Revenue, and Trade Offices
District	Tigray	Tigray Hintalo Wajerat Agriculture, Education, Health, Revenue, and Trade Offices
District	Tigray	Tigray Tahtay Koraro Agriculture, Education, Health, Revenue, and Trade Offices
District	Tigray	Tigray Wereilehi Agriculture, Education, Health, Revenue, and Trade Offices

Table A2: Description of Survey and Administrative Data

Summary statistics

	Source	(1) Mean	(2) Standard deviation	(3) Skewness	(4) Autogregressive parameter from DF regression, no time trends	(5) Autogregressive parameter from DF regression, linear time trends	(6) Z Statistic from Harris Tsavalis Unit Root Test (linear time trends)	(7) P-value from Harris Tsavalis Unit Root Test (linear time trends)	(8) Fixed-N exact critical value for the Im-Pesaran-Shin test at the 5% level (linear time trends)	(9) Fixed-N Im-Pesaran-Shin Test statistic	(10) N
Demographic and Socio-economic Characteristics											
What do you think is recorded as the population of [INSERT WOREDA] according to official administrative data in the Census of 2007 Gregorian Calendaar (1999/2000 Ethiopian Calendar)?	Census 2007	99,605	61,868	1.19							740
What do you think is recorded as the percentage of rural inhabitants of [INSERT WOREDA] according to official administrative data in the Census of 2007 Gregorian Calendar (1999/2000 Ethiopian Calendar)?	Census 2007	0.84	0.26	-2.66							726
What do you think is recorded as the unemployment rate of [INSERT WOREDA] according to official administrative data in the Census of 2007 Gregorian Calendar (1999/2000 Ethiopian Calendar)?	Census 2007	4.07	5.63	2.21							740
Education indicators											
What do you think the primary (grades 1-8) enrolment numbers are for [INSERT WOREDA] according to official administrative data?	EMIS 2006/07-2013/14	21,929	14,020	1.06	0.61	0.05	-12.36	0.00	-2.34	-2.63	705
What do you think the primary (grades 1-8) pupil-section ratio is for [INSERT WOREDA] according to official administrative data?	EMIS 2006/07-2013/14	57.17	26.71	5.32	0.41	-0.03					705
What do you think the primary (grades 1-8) pupils-per-school ratio is for [INSERT WOREDA] according to official administrative data?	EMIS 2006/07-2013/14	550	229	1.61	0.57	0.16	-5.46	0.00	-2.34	-2.79	705
What do you think the primary (grades 1-8) pupil-teacher ratio is for [INSERT WOREDA] according to official administrative data?	EMIS 2006/07-2013/14	97.61	354	15.1	-0.15	0.02					467
Health indicators											
What do you think the Proportion of pregnant women who attended ANC4+ during the current pregnancy is for [INSERT WOREDA] according to official administrative data (the baseline of the 2007 Core Plan)?	HSDP Core Plan 2008/09-2013/14	0.51	0.29	0.08	-0.13	-0.53	-6.81	0.00			688
What do you think the Contraceptive Acceptance Rate is for [INSERT WOREDA] according to official administrative data (the baseline of the 2007 Core Plan)?	HSDP Core Plan 2008/09-2013/14	0.59	0.29	-0.23	-0.19	-0.52	-6.38	0.00			688
What do you think the Rate of Births Attended by Skilled Health Personnel is for [INSERT WOREDA] district government according to official administrative data (the baseline of the 2007	HSDP Core Plan 2008/09-2013/14	0.33	0.22	2.62	-0.19	-0.60	-8.26	0.00			688
What do you think the Proportion of infants fully immunized is for [INSERT WOREDA] according to official administrative data (the baseline of the 2007 Core Plan)?	HSDP Core Plan 2008/09-2013/14	0.74	0.22	-1.24	-0.08	-0.57	-8.06	0.00			688
Agriculture indicators											
How much land do you think is used for agricultural purposes for [INSERT WOREDA] district government according to data from the WCBS?	WCBS Round III (2009/10) & Round V (2012/13)	29,366	33,892	6.49	1.01	1.01					203
How much land do you think is used for pastoral purposes for [INSERT WOREDA] according to data from the WCBS?	WCBS Round III (2009/10) & Round V (2012/13)	27,009	75,005	5.97	1.00	1.00					195
What do you think is the agricultural income per household per year in [INSERT WOREDA] according to data from the WCBS?	WCBS Round III (2009/10) & Round V (2012/13)	11,738	15,868	4.58	0.32	0.32					146
What do you think is the share of households dependent on subsistence agriculture in [INSERT WOREDA] according to data from the WCBS?	WCBS Round III (2009/10) & Round V (2012/13)	24.29	23.23	1.56	0.92	0.92					132
Revenue indicators											
How many tax identification numbers do you think were issued in the last financial year according to data from the WCBS for [INSERT WOREDA]?	WCBS Round III (2009/10) & Round V (2012/13)	425	478	2.42	0.33	0.33					146
What proportion of income for this district government do you think is received from regional recurrent block grants according to data from the WCBS for [INSERT WOREDA]?	WCBS Round III (2009/10) & Round V (2012/13)	0.81	0.18	-1.97	-0.22	-0.22					190
What proportion of income for this district government do you think is received from own sources according to data from the WCBS for [INSERT WOREDA]?	WCBS Round III (2009/10) & Round V (2012/13)	0.20	0.15	2.46	-0.08	-0.08					177
Trade indicators											
How many business licenses do you think were issued in the last financial year according to data from the WCBS for [INSERT WOREDA]?	WCBS Round III (2009/10) & Round V (2012/13)	376	481	5.19	1.36	1.36					177
What do you think is the total revenue collected from issuing and renewing business licenses according to data from the WCBS for [INSERT WOREDA]?	WCBS Round III (2009/10) & Round V (2012/13)	108,473	218,841	8.69	0.32	0.32					172
Civil Service Rules											
What is the amount of regular working hours for a civil servants according to the Civil Service Proclamation? (Paragraph 32)	Civil Servants Proclamation	39									
What is the allowed amount of annual leave in the first year of service according to the Civil Service Proclamation? (Paragraph 37, Item 1)	Civil Servants Proclamation	20									
What is the maximum amount of annual leave for a civil servant who has served for more than one year according to the Civil Service Proclamation? (Paragraph 37, Item 2)	Civil Servants Proclamation	30									
What is the allowed amount of maternity leave according to the Civil Service Proclamation? (Paragraph 41, Item 2)	Civil Servants Proclamation	90									
What are the types of different disciplinary actions that a civil servant can face for the breach of discipline according to the Civil Service Proclamation? (Paragraph 67)	Civil Servants Proclamation	6									
What is the notice time required before resigning according to the Civil Service Proclamation? (Paragraph 78, Item 1)	Civil Servants Proclamation	30									

Notes: The unit of observation is the distict. EMIS is the Education Management Information System of the Ministry of Education, Ethiopia. HSDP is the Health Sector Development Plan of the Ministry of Health, Ethiopia. WCBS is the District government and City Benchmarking Survey, undertaken by the Ministry of Public Service and Human Resource Development and World Bank. Skewness refers to Pearson's moment coefficient of skewness. Column 4 reports the autogregressive coefficient from the standard Dickey Fuller regression in a panel / time-series dataset, controlling for district-specific fixed effects. Column 5 reports the autogregressive coefficient from the Dickey Fuller regression in a panel / time-series dataset, controlling for district-specific fixed effects and district-specific linear time trends. Column 6 reports the test statistics from the Harris Tsavaliz test, which tests for the existence of a unit-root in a panel dataset, allowing for district-specific fixed effects and linear district-specific time trends; the model assumes a common autogregressive parameter for each series; assumes that the number of districts tends to infinity while the total number of time periods is fixed and the errors are serially uncorrelated; the test can only be conducted on a balanced panel dataset; the null hypothesis is that there exists a (common) unit root. Column 7 reports the P-value from the Harris Tsavalis test, corresponding to the Z-Statistic reports in Column 5. Column 8 reports the fixed-N critical value for the Im_Pesaran-Shin Test of a unit root and Column 9 reports the test statistic for the series, allowing for district-specific fixed effects and district-specific linear time trends, and a district-specific autoregressive parameter. The Im-Pesaran-Shin test allows for for panel-specific autoregressive parameters and an unblanaced panel dataset; the Dickey Fuller regression is fitted to each panel separately and an average test statistic is used, under the assumption that the errors are serially uncorrelated, with fixed N and fixed T; the null hypothesis is that all of the series posesses a unit root, against the alternative that some of these series contain unit roots. Test statistics for the pupil-class ratio and pupil-teacher ratio are not displayed due to gaps in the panel series; test statistics and autoregressive

Table A3: World Management Survey Indicators

Aggregate Index	Topic	Question	Score 1	Score 3	Score 5
Monitoring	Monitoring	In what kind of ways does your Directorate track how well it is delivering services? Can you give me an example?	Directorate does not track performance.	Directorate tracks a number of performance indicators. These are seen and reviewed by senior management only.	Full set of indicators are tracked formally and continuously. Reviews are conducted regularly and involve representative of all directorate staff groups. The results of the review are formally communicate to all directorate staff.
Monitoring	Monitoring	Are you involved in performance review for your Directorate? If so, how often does this occur?	Not involved in performance review;	Bi-annually	Monthly
Other	Targeting	Does your Directorate have a clear set of targets derived from the organization's goals and objectives? Are they used to determine your work schedule?	The directorate does not have defined targets.	Targets are assigned to the directorate, as well as to the manager and employee levels, and these are generally well understood by mid-level staff. However the tasks assigned to staff are not always related to those targets.	Targets are clearly defined for the directorate, manager, and employee levels, and are well understood by all staff. All tasks are directly derived from the targets, which are regularly reviewed to ensure they remain on track.
Other	Targeting	When you arrive at work each day, do you and your colleagues know what their individual roles and responsibilities are in achieving the organization's goals?	Staff do not know what their roles and responsibilities are.	Staff have a good idea of their roles and responsibilities but it is not always clear how they contribute to their organization's goals.	Staff have a very good understanding of their roles and responsibilities. Their own roles and goals are clearly interconnected to those of their organization.
Other	Targeting	How are targets and performance measures communicated to staff in your directorate?	Neither targets nor performance measures are communicated to staff.	Targets and performance measures are formally communicated to managers and team leaders.	Targets and performance measures are formally communicated and understood by all staff.
Other	Performance incentives	How would under-performance be tolerated in your Directorate? Can you give me an example of how such a case would be dealt with?	Poor performers stay in their positions (no consequences).	Poor performance Is identified through evaluation and is addressed through concrete action. Although this applies to most staff, some individuals/staff groups get away with it.	Poor performers are identified through regular reviews and are put on a formal performance improvement plan immediately. This applies to all staff.
Other	Performance incentives	Given past experience, have members of [respondent's organization] been disciplined for breaking the rules of the civil service?	There are no consequences for bad behaviour/ breaking the rules.	Bad behaviour is addressed through concrete action, but the underlying issues are not addressed.	Bad behaviour/ breaking the rules is addressed through concrete action. If any employee breaks the rules, the underlying issues will be identified and rectified. This applies to all employees.
Other	Performance incentives	Does your Directorate use performance, targets, or indicators for tracking and rewarding (financially or non-financially) the performance of its employees?	Staff are rewarded equally (or not rewarded) irrespective of performance. Individual performance is not tracked formally	There is a formal staff evaluation system in place and performance is rewarded (financially or non-financially). However, there are no clear system or criteria for rewarding staff	There is a formal staff evaluation system and performance is rewarded (financially or non- financially). Rewards are given as a consequence of well-defined and monitored individual achievements. This applied to all staff.
Other	Roles	When staff in your Directorate are given tasks in their daily work, how much discretion do they have to carry out their assignments? Can you give me an example?	How officers carry out their assignments is decided by senior managers. Officers have no say.	How officers carry out their assignments is jointly decided by the officer and senior managers. Senior managers tend to drive the decisions.	Officers have complete autonomy in deciding how to carry out their tasks.
Other	Roles	Can most staff in your Directorate make substantive contributions to the policy formulation and implementation process?	Staff do not contribute to policy formulation, nor to decisions about implementation.	Staff can contribute to policy formulation and decisions about implementation, but there is no formal forum through which to do this. Contributions typically only occur when problems arise.	Management expects all staff to contribute to policy formulation and decisions about implementation (formally or informally), and considers this part of their duties.
Other	Roles	Is the workload of achieving your Directorate's targets evenly distributed across its different employees, or do some groups consistently shoulder a greater burden than others?	A small minority of staff undertake the vast majority of work within the directorate.	The burden of the directorate's work is more or less distributed equally among staff. A small minority get away with working significantly less than others.	The burden of the directorate's work is distributed equally among staff. Tasks are assigned in such a way that the amount of time required and the level of difficulty are balanced out so no member of staff finds him/herself overburdened.
Other	Roles	Thinking about all the projects that your Directorate has been involved in since your appointment here, would you say that managers and supervisors try to use the right staff for the right job?	Staff are allocated to tasks randomly.	Managers try to use the right staff for the right job but do not go to great lengths to ensure this, or are met with institutional constraints which may prevent them from doing so.	The right staff are always used for a task. Allocation of tasks is based on staffs' documented skills and competencies.
Other	Flexibility	Does your Directorate make efforts to adjust to the specific needs and specific requirements of communities, clients, or other stakeholders?	The directorate uses the same procedures no matter what.	The directorate tailors procedures to the specific needs of its stakeholders, but struggles when those needs are complex.	The directorate tailors all procedures to the specific needs of its stakeholders. The evolution of those needs results in adaptation to plans, project and policies.
Other	Flexibility	How flexible would you say your Directorate is in terms of responding to new and improved work practices or reforms?	New practices are not adopted/ integrated in the directorate.	New ideas or practices are adopted, but in an informal and/ or isolated manner. The directorate encourages the adoption of new practices, however it is slow to integrate them into its operations (more than a year).	The adoption of new ideas and practices is an integral part of the directorate's work. New practices are regularly reviewed and considered, and once adopted and integrated across the directorate within 6 months.
Other	Staff involvement/ contribution	How do problems in your directorate get exposed and fixed?	Ad-hoc, no set process for improvement Deal with problems as they arise without following an established procedureOnce fixed, no further action taken No suggestions from staff	Existing process to deal with problems Improvements made through meetings Focus on finding solutions, not prevention of future problems Suggestions from staff involved through meetings (formal or informal)	Exposing problems and suggesting solutions and improvements is part of all staffs' daily duty. Continuous improvement is part of the culture of the organization.
Other	Staff involvement/ contribution	What kind of feedback do you get in staff meetings?	No feedback from staff.	Staff provide feedback in meetings but in an unstructured manner. Focus on bad performance.	Staff provide the feedback on which action plans will be based. Focus on both good and bad performance. Details of the meetings are recorded and communicated to all staff.
Other	Staff involvement/ contribution	Let's say you've agreed to a follow up plan at one of your meetings, what would happen if the plan wasn't enacted?	No action taken. No changes made in the operations process.	Failure can be found in regular meetings (weekly, even monthly for long-term plans) or at standard points before the deadline. Plans can be altered in order to achieve expected results on time.	In addition to 4, tools can be checked up and reported to the manager in charge. Meetings (formal/ informal) are held to look into the root causes of problems and preventive actions are taken for future similar task.
Other	Staffing	Do you think the management of your Directorate think about attracting talented people to your Directorate and then doing their best to keep them? For example, by ensuring they are happy and engaged with their work.	Directorate does not put emphasis on talent	Senior management believes that attracting and developing talent is important, but there is no clear system for identifying, attracting or retaining such talent.	Senior management believes that attracting and developing talent is important. There is a clear system for identifying and attracting talent, developing and retaining talent.
Other	Staffing	If two senior level staff joined your Directorate five years ago and one was much better at their work than the other, would he/she be promoted through the service faster?	No promotion system (no one in the organization has been promoted for years) The promotion system is based on tenure	The promotion system is based on performance. Organization may have internal limitations (e.g. few position openings), but do everything to get around them (e.g. extra training).	Promotion system is based on performance. Organization actively identifies, develops and promotes top performers. Regular assessments, clear set of indicators and personalised career plans for individuals (regularly revised).

Table 31: Information, Hierarchy, Marginal Costs and Marginal Benefits

Table A4: Determinants of Superior Information in Hierarchies

Dependent Variable: Z-score of errors across policy domains

Analysis of Variance (ANOVA) Estimates in column (1); OLS estimates in column (2)

Partial sum of squares and F-statistic p-values in parentheses in column (1); OLS coefficient and clustered standard error at the district-government-sector level, in parentheses in column (2); F-statistics and associated p-values from the joint test across all categories for factor variables preceded by "F=" and "P=" respectively in column (2)

	(1) ANOVA	(2) Regression coefficients
Hierarchical structure		
District Government [Yes=1]	8.71	-0.33**
	[0.00]	[0.14]
Manager	0.04	0.012
	[0.83]	[0.072]
Sector	3.61	*F=0.6*
	[0.41]	*P=0.67*
Marginal costs		
Feature of Claim: Indicator	27.09	*F=1.92*
	[0.05]	*P=0.02*
Feature of Claim: Percentile	9.38	*F=2.73*
	[0.01]	*P=0.07*
Bureaucrat Characteristic: Years in position	0.00	-0.00049
	[0.95]	[0.012]
Bureaucrat Characteristic: Years in organization	0.19	-0.0022
	[0.65]	[0.0047]
Bureaucrat Characteristic: Years in civil service	0.00	0.00013
	[0.98]	[0.0040]
Bureaucrat Characteristic: Undergraduate education [degree=1]	0.52	-0.069
	[0.45]	[0.12]
Bureaucrat Characteristic: Postgraduate education [masters=1]	0.02	-0.0092
	[0.89]	[0.078]
Bureaucrat Characteristic: Gender [female=1]	0.53	0.046
	[0.44]	[0.065]
Organization Characteristic: MIS exists	0.04	0.011
	[0.84]	[0.18]
Organization Characteristic: Proportion of projects civil servant has adequate information on service delivery	11.74	0.014***
	[0.00]	[0.0052]
Organization Characteristic: Proportion of projects information flows effectively within organizations	4.84	-0.012**
	[0.02]	[0.0057]
Organization Characteristic: Proportion of projects information flows effectively across organizations	0.08	0.0011
	[0.77]	[0.0055]
Organization Characteristic: Span of control	3.85	-0.0090*
	[0.04]	[0.0053]
Organization Characteristic: Number of principals	3.44	0.075
	[0.05]	[0.050]
Marginal benefits		
Management practices: Monitoring	4.35	-0.15**
	[0.03]	[0.072]
Management practices: Other	3.33	0.16
	[0.05]	[0.11]
Alignment: Mission alignment index	1.10	0.14
	[0.27]	[0.18]
Model	188.47	
	[0.00]	
Residual	1239.29	
District fixed effects	Yes	
Adjusted R-squared	-	0.080
Number of observations [clusters]	1455	1455 [92]

Notes: *** denotes significance at 1%, ** at 5%, and * at 10% level. Column (1) reports ANOVA partial sum of square estimates. Column (2) reports OLS regression coefficients. Standard errors clustered at the district-government-sector level. Column (2) reports the standard error in brackets; for factor variables with mulitple categories, the F-statistic of joint test of all category indicators are reported after F= and the p-value is reported below after P=. The unit of analysis in all columns is the claim. The sample used for the analysis is the set of respondents that make claims about the same districts. Sector fixed effects are indicators for whether the claim is within an agriculture, education, health, revenue, or trade organization. Percentile fixed effects are indicators for the 75th, 50th or 25th percentile for the respective indicator for the respective region. Indicator fixed effects are indicator variables for which specific variable the claim is being made over. Manager is equal to one if the claim is from a manager. MIS exists is the organization proportion of managers that respond "Yes" to "Do you currently have a Management Information System (MIS) programme in place and functioning?". Proportion of projects civil servant has adequate information on service delivery is the organization average response to "In what proportion of projects would you say you have necessary information regarding the current state of service delivery in your jurisdiction / area of work?". Proportion of projects information flows effectively within organizations is the organization average response to "In what proportion of projects would you say information flows effectively between individuals in the same organization?". Proportion of projects information flows effectively across organizations is the organization average response to "In what proportion of projects would you say Information flows effectively between organizations?". Span of control is the organizations average response across managers of the response to "How many personnel do you manage?". Number of principals is the organization average response of employees to "How many people would you say regularly give you tasks as part of your formal work duties?". Management practice: Monitoring is the organization average of the monitoring topic, construced from the z-scores of each individual item noted in table A3. Management practice: Other is the organization average of the average across the z-scores of targeting, incentives, roles, flexibility, staffing, and staff involvement, detailed in table A3. Alignment: Mission alignment index is the z-score of the row mean of the organization-average response to "To what extent do you believe that your organization's mission is aligned to your own mission?" and the organization proportion of employees selecting "I am contributing to fulfilling that mission on an everyday basis" to "How consistent with your mission are the various tasks and activities assigned to you on a day-to-day basis?". Figures are rounded to two decimal places or two significant figures.

Table A5: Decentralization and Information Robustness

Dependent Variable: Z-score of errors across policy domains in columns (1) to (5) and column (8); Rank of errors, in terms of the proportion within distribution, across policy domains in column (6); Skewness adjusted z-scores of errors across policy domains in column (7)

OLS Estimates

Standard Errors: Clustered at the district government-sector level in column (1) and columns (6) to (8); Clustered at the claimed district government level in column (2); Clustered at the respondent level in column (3); Clustered at the organization-of-respondent level in column (4); Clustered at the region-sector level in column (5).

	(1) Disaggregating central dummy	(2) Clustering at claimed district government level	(3) Clustering at respondent level	(4) Clustering at the organization-of-respondent level	(5) Clustering at region-sector level	(6) Rank of Error (proportion)	(7) Skewness adjusted z-score	(8) Management controls
District Government [Yes=1]	-0.31***	-0.33***	-0.33***	-0.33***	-0.33***	-0.10***	-0.39***	-0.28*
	[0.098]	[0.096]	[0.064]	[0.081]	[0.094]	[0.027]	[0.098]	[0.16]
Federal Government [Yes=1]	0.042							
	[0.091]							
Manager	0.039	0.034	0.034	0.034	0.034	-0.00034	0.029	0.060
	[0.067]	[0.070]	[0.065]	[0.068]	[0.073]	[0.018]	[0.059]	[0.11]
Sector fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
District government fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No
Indicator fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Percentile fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Individual controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Region fixed effects	No	No	No	No	No	No	No	Yes
Marginal costs and noise controls	No	No	No	No	No	No	No	Yes
Adjusted R-squared	0.068	0.069	0.069	0.069	0.069	0.11	0.11	0.057
Observations [clusters]	1455 [92]	1455 [42]	1455 [687]	1455 [122]	1455 [38]	1455 [92]	1455 [92]	1455 [92]
Federal respondents [percentage]	195 [28%]	195 [28%]	195 [28%]	195 [28%]	195 [28%]	195 [28%]	195 [28%]	195 [28%]
Regional respondents [percentage]	234 [34%]	234 [34%]	234 [34%]	234 [34%]	234 [34%]	234 [34%]	234 [34%]	234 [34%]
District government respondents [percenta	260 [38%]	260 [38%]	260 [38%]	260 [38%]	260 [38%]	260 [38%]	260 [38%]	260 [38%]

Notes: *** denotes significance at 1%, ** at 5%, and * at 10% level. Clustered standard errors are in parentheses. Standard errors clustered at the district government-sector interacted level in column (1) and columns (6) to (8); clustered at the claimed district level in column (2); clustered at the respondent level in column (3); clustered at the organization-of-respondent level in column (4); clustered at the region-sector interacted level in column (5). All columns report OLS estimates. The unit of analysis in all columns is the claim. The sample used for the analysis is the set of respondents that make claims about the same districts. The dependent variable is the z-score of absolute error across all policy domain items in columns (1) to (5) and column (6), where the z-score is calculated within the distribution of errors within an indicator; the rank of the error in proportion terms (proportion of the maximum rank within the indicator) within the distribution of errors within an indicator in column (6); the skewness-adjusted z-score is the dependent variable in column (7) - the skewness adjusted z-score is the z-score of the absolute error if the variable is a proportion variable (between 0 and 1), if the variable is not a proportion and is statistically significantly skewed relative to the normal distribution (Royston,1992), then the log of the variable is taken and then the z-score is created from the logged values. Sector fixed effects are indicators for whether the claim is within an agriculture, education, health, revenue, or trade organization. Percentile fixed effects are indicators for the 75th, 50th or 25th percentile for the respective indicator for the respective region. District government fixed effects identify the district that the respondent is making claims about. Indicator fixed effects control for the service delivery indicator within which the claim is made. Management is equal to one if the claim is from a director. Individual characteristics included in the specification: tenure in the civil service of the respondent making the claim; tenure in the position; tenure in the organization; an indicator for whether the respondent is female; an indicator for whether the respondent holds an undergraduate degree; an indicator for whether the respondent holds a masters degree. Marginal costs include the proportion of managers in the organization that state that a management information system is in place; the organization average response to "In what proportion of projects would you say information flows effectively between individuals in the same organization?"; the organization average response to "In what proportion of projects would you say Information flows effectively between organizations?"; the organization average response of managers to "How many personnel do you manage?"; the organization average response of employees to "How many people would you say regularly give you tasks as part of your formal work duties?"; alignment (the z-score of the row mean of the organization-average response to "To what extent do you believe that your organization's mission is aligned to your own mission?" and the organization proportion of selecting "I am contributing to fulfilling that mission on an everyday basis" to "How consistent with your mission are the various tasks and activities assigned to you on a day-to-day basis?"). Noise controls include the time of day of the survey, day-of-survey fixed effects, enumerator fixed effects, an indicator for the enumerator's subjective assessment of the quality of the interview; an indicator for the duration of the interview (decile of duration fixed effects); controls for the average tenure of the managers in the organization and the proportion of female managers in the organization. Figures are rounded to two significant figures.

Table A6: Heterogeneity in Effects: Quantile Regression By Absolute Error

Dependent Variable: Z-score of errors across policy domains in columns

Quantile regression estimates

Standard Errors: Clustered at the district government-sector level

	(1) Percentile = 0.1	(2) Percentile = 0.2	(3) Percentile = 0.3	(4) Percentile = 0.4	(5) Percentile = 0.5	(6) Percentile = 0.6	(7) Percentile = 0.7	(8) Percentile = 0.8	(9) Percentile = 0.9
Federal Government [Yes=1]	0.0095	0.045	0.057	0.055	0.071	0.014	-0.011	-0.021	0.066
	[0.034]	[0.038]	[0.053]	[0.055]	[0.072]	[0.078]	[0.092]	[0.10]	[0.13]
District Government [Yes=1]	-0.076*	-0.15***	-0.21***	-0.21***	-0.20***	-0.21**	-0.27**	-0.40***	-0.26*
	[0.039]	[0.053]	[0.080]	[0.078]	[0.076]	[0.088]	[0.10]	[0.11]	[0.14]
Manager [Yes=1]	0.019	0.019	0.031	0.038	0.0074	-0.014	0.0035	0.023	0.083
	[0.029]	[0.031]	[0.031]	[0.032]	[0.051]	[0.066]	[0.075]	[0.090]	[0.10]
Sector fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
District government fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Indicator fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Percentile fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Individual controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations					1455 [92]				

Notes: *** denotes significance at 1%, ** at 5%, and * at 10% level. Standard errors are in parentheses. Standard errors clustered at the district-government-sector interacted level. Columns report quantile regression estimates for different quantiles of the distribution of the z-score of absolute errors. The unit of analysis in all columns is the claim. The sample used for the analysis is the set of respondents that make claims about the same districts. The dependent variable is the z-score of the absolute error. Sector fixed effects are indicators for whether the claim is within an agriculture, education, health, revenue, or trade organization. Percentile fixed effects are indicators for the 75th, 50th or 25th percentile for the respective indicator for the respective region. Indicator fixed effects control for the service delivery indicator within which the claim is made. Manager is equal to one if the claim is from a manager. Individual characteristics included in the specification: tenure in the civil service of the respondent making the claim; tenure in the organization; an indicator for whether the respondent is female; an indicator for whether the respondent holds an undergraduate degree; an indicator for whether the respondent holds a masters degree. Figures are rounded to two decimal places or two significant figures.

Table A7: Decentralisation and Rules

Dependent Variable: Z-score of errors across civil service rules in columns (1) to (9); Rank of errors across civil service rules in column (10); Skewness-adjusted z-score of errors across civil service rules in column (11)

OLS Estimates

Standard Errors: Clustered by organization-of-respondent in columns (1) to (7) and columns (10) to (11); Clustered at the respondent level in column (8); Clustered at the region and sector level in column (9)

	(1) Unconditional	(2) Sector FEs	(3) District government FEs	(4) Indicator FEs	(5) Percentile FEs	(6) Management	(7) Individual Characteristics	(8) Clustering at respondent level	(9) Clustering at region-sector level	(10) Rank of Error (proportion)	(11) Skewness adjusted z-score
Federal Government	-0.14***	-0.14***	-0.15***	-0.15***	-0.15***	-0.14***	-0.088**	-0.088***	-0.088**	-0.037***	-0.13***
	[0.042]	[0.040]	[0.035]	[0.035]	[0.034]	[0.035]	[0.035]	[0.033]	[0.036]	[0.010]	[0.036]
District government	-0.0057	-0.00058	-0.011	-0.011	-0.012	-0.011	0.015	0.015	0.015	-0.014	0.061
	[0.042]	[0.042]	[0.039]	[0.039]	[0.038]	[0.038]	[0.039]	[0.037]	[0.039]	[0.0099]	[0.038]
Management [Yes=1]						0.047	0.00088	0.00088	0.00088	0.011	-0.0084
						[0.033]	[0.030]	[0.032]	[0.031]	[0.0072]	[0.028]
Sector fixed effects	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
District government fixed effects	No	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Indicator fixed effects	No	No	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Percentile fixed effects	No	No	No	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Individual controls	No	No	No	No	No	No	Yes	Yes	Yes	Yes	Yes
Adjusted R-squared	0.0035	0.0042	0.012	0.011	0.011	0.011	0.032	0.032	0.032	0.048	0.40
Observations [clusters]	6201 [122]	6201 [122]	6201 [122]	6201 [122]	6201 [122]	6201 [122]	6201 [122]	6201 [689]	6201 [38]	6201 [122]	6201 [122]
Federal respondents [percentage]	195 [28%]	195 [28%]	195 [28%]	195 [28%]	195 [28%]	195 [28%]	195 [28%]	195 [28%]	195 [28%]	195 [28%]	195 [28%]
Regional respondents [percentage]	234 [34%]	234 [34%]	234 [34%]	234 [34%]	234 [34%]	234 [34%]	234 [34%]	234 [34%]	234 [34%]	234 [34%]	234 [34%]
District government respondents [percentage]	260 [38%]	260 [38%]	260 [38%]	260 [38%]	260 [38%]	260 [38%]	260 [38%]	260 [38%]	260 [38%]	260 [38%]	260 [38%]

Notes: *** denotes significance at 1%, ** at 5%, and * at 10% level. Clustered standard errors are in parentheses. Standard errors clustered at the organization level unless stated otherwise. All columns report OLS estimates. The unit of analysis in all columns is the claim. The sample used for the analysis is the set of respondents that make claims about the same districts. The dependent variable is the mean z-score of absolute errors across all policy domain items claimed by the same respondent about the same district. Sector fixed effects are indicators for whether the claim is within an agriculture, education, health, revenue, or trade organization. Percentile fixed effects are indicators for the 75th, 50th or 25th percentile for the respective indicator for the respective region. District government fixed effects identify the district that the respondent is making claims about. Indicator fixed effects control for the service delivery indicator within which the claim is made. Management is equal to one if the claim is from a manager. Individual characteristics included in the specification: tenure in the civil service of the respondent making the claim; tenure in the position; tenure in the organization; an indicator for whether the respondent is female; an indicator for whether the respondent holds an undergraduate degree; an indicator for whether the respondent holds a masters degree. Column 8 clusters the errors at the respondent level. Column 9 clusters the error at the region-sector level where the region is that in which the district being referred to is situated in. Column 10 runs the preferred specification of the model, as per column 7, clustering at the organization levell, but with the dependent variable as the rank of the absolute error across all policy domains; the rank is represented as the proportion within the distrbution of the errors within the indicator. Column 12 runs the preferred specification of the model, as per column 7, clustering at the organization level, but with the dependent variable as the skewness adjusted z-score of the absolute error across all policy domains; the skewness adjusted z-score is simply the z-score if the variable is a proportion; if the variable is not a proportion and is statistically significantly skewed relative to the normal distribution (Royston,1991), then the log of the variable is taken and then the z-score is created from the logged values. Figures are rounded to two significant figures.

Table A8: Decentralization and Management

Dependent Variable: Management practice: Monitoring in columns (1) to (4); Management practice: Other in columns (5) to (8)
OLS Estimates
Robust standard errors

	Management practice: Monitoring				Management practice: Other			
	(1) Unconditional	(2) Sector FE	(3) Individual characteristics (organization average)	(4) Marginal costs	(5) Unconditional	(6) Sector FE	(7) Individual characteristics (organization average)	(8) Marginal costs
District Government [Yes=1]	0.13	0.16	0.41**	0.54**	-0.12	-0.10	0.14	0.25*
	[0.15]	[0.15]	[0.21]	[0.21]	[0.12]	[0.12]	[0.14]	[0.14]
Sector fixed effects	No	Yes	Yes	Yes	No	Yes	Yes	Yes
Individual characteristics (organization average)	No	No	Yes	Yes	No	No	Yes	Yes
Marginal costs	No	No	No	Yes	No	No	No	Yes
Adjusted R-squared	-0.0023	0.034	0.078	0.25	0.00018	0.049	0.10	0.24
Observations	122	122	122	122	122	122	122	122

Notes: *** denotes significance at 1%, ** at 5%, and * at 10% level. Standard errors are in parentheses. Robust standard errors. All columns report OLS estimates. The unit of analysis in all columns is the organization. The tiers of government are grouped into District Government and Centralized Government which incorporates both Federal and Regional tiers, hence the omitted category for tier in the regressions is Centralized Government. The sample used for the analysis is the set of respondents that make claims about the same districts. The dependent variable is Management practice: Monitoring in columns (1) to (4) and Management practice: Other in columns (5) to (8). Management practice: Monitoring is the organization average of the monitoring topic, construced from the z-scores of each individual item noted in table A3. Management practice: Other is the organization average of the average across the z-scores of targeting, incentives, roles, flexibility, staffing, and staff involvement, detailed in table A3. Management practices are in z-scores. Individual characteristics (organization average) include the organization average of: the number of years in current position; the number of years in current organization; the number of years in the service; an indicator for whether the employee has an undergraduate degree; an indicator for whether the employee has a masters degree; and an indicator for whether the employee is female. Marginal costs include the proportion of managers in the organization that state that a management information system is in place; the organization average response to "In what proportion of projects would you say information flows effectively between individuals in the same organization?"; the organization average response to "In what proportion of projects would you say Information flows effectively between organizations?"; the organization average response of managers to "How many personnel do you manage?"; the organization average response of employees to "How many people would you say regularly give you tasks as part of your formal work duties?"; alignment (the z-score of the row mean of the organization-average response to "To what extent do you believe that your organization's mission is aligned to your own mission?" and the organization proportion of selecting "I am contributing to fulfilling that mission on an everyday basis" to "How consistent with your mission are the various tasks and activities assigned to you on a day-to-day basis?"). Figures are rounded to two significant figures.

Table A9: Errors and Appraisal Scores

Dependent Variable: Within-organization z-score of civil servant's total appraisal score

OLS Estimates

Standard Errors: Clustered at the district government-sector level

	(1) Unconditional	(2) Tier fixed effects	(3) Sector fixed effects	(4) Indicator fixed effects	(5) Percentile fixed effects	(6) Region fixed effects	(7) Individual characteristics	(8) Marginal costs	(9) Management practices
Absolute error (z-score)	-0.021	-0.023*	-0.023*	-0.024*	-0.024*	-0.028**	-0.021*	-0.026**	-0.024*
	[0.013]	[0.013]	[0.013]	[0.013]	[0.013]	[0.013]	[0.012]	[0.012]	[0.012]
Don't know	-0.085	-0.12	-0.13	-0.14	-0.14	-0.096	-0.083	-0.031	-0.040
	[0.066]	[0.080]	[0.086]	[0.088]	[0.088]	[0.10]	[0.092]	[0.094]	[0.092]
Tier fixed effects	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Sector fixed effects	No	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Indicator fixed effects	No	No	No	Yes	Yes	Yes	Yes	Yes	Yes
Percentile fixed effects	No	No	No	No	Yes	Yes	Yes	Yes	Yes
Region fixed effects	No	No	No	No	No	Yes	Yes	Yes	Yes
Individual characteristics	No	No	No	No	No	No	Yes	Yes	Yes
Marginal costs and noise controls	No	No	No	No	No	No	No	Yes	Yes
Management practies	No	No	No	No	No	No	No	No	Yes
Adjusted R-squared	0.0024	0.0033	0.0035	-0.00049	-0.00099	0.018	0.082	0.14	0.14
Observations [clusters]	3963 [135]	3963 [135]	3963 [135]	3963 [135]	3963 [135]	3963 [135]	3963 [135]	3963 [135]	3963 [135]

Notes: *** denotes significance at 1%, ** at 5%, and * at 10% level. Clustered standard errors are in parentheses. Standard errors clustered at the district government-sector interacted level. All columns report OLS estimates. The unit of analysis in all columns is the claim of the district government-level official. The sample used for the analysis is the set of respondents making claims for which there is also evaluaton data for the civil servant. The dependent variable is the within-organization z-score of the civil servant's evaluation outcome as per their annual appraisal. Tier fixed effects control for whether the respondent works in a federal, regional, or district government office. Sector fixed effects are indicators for whether the claim is within an agriculture, education, health, revenue, or trade organization. Indicator fixed effects control for the specific variable being claimed over. Individual characteristics includes: years in position; years in the organization; years in the civil service; an indicator for whether the respondent has an undergraduate degree; and an indicator for whether the respondent has a masters degree; an indicator for whether the respondent is female; and an indicator for whether the respondent is a manager. Marginal costs include the proportion of managers in the organization that state that a management information system is in place; the organization average response to "In what proportion of projects would you say information flows effectively between individuals in the same organization?"; the organization average response to "In what proportion of projects would you say Information flows effectively between organizations?"; the organization average response of managers to "How many personnel do you manage?"; the organization average response of employees to "How many people would you say regularly give you tasks as part of your formal work duties?"; alignment (the z-score of the row mean of the organization-average response to "To what extent do you believe that your organization's mission is aligned to your own mission?" and the organization proportion of selecting "I am contributing to fulfilling that mission on an everyday basis" to "How consistent with your mission are the various tasks and activities assigned to you on a day-to-day basis?"). Noise controls include the time of day of the survey, day-of-survey fixed effects, enumerator fixed effects, an indicator for the enumerator's subjective assessment of the quality of the interview; an indicator for the duration of the interview (decile of duration fixed effects); controls for the average tenure of the managers in the organization and the proportion of female managers in the organization. Management practices incude Management practice: Monitoring and Managament practice: Other. Management practice: Monitoring is the organization average of the monitoring topic, construced from the z-scores of each individual item noted in table A3. Management practice: Other is the organization average of the average across the z-scores of targeting, incentives, roles, flexibility, staffing, and staff involvement, detailed in table A3. Figures are rounded to two significant figures.

Table A10: Errors and Service Delivery

Dependent Variable: Z-Score of service delivery indicator
OLS Estimates
Standard Errors: Clustered at the district government-sector level

	(1) Unconditional	(2) Sector FE	(3) Indicator FE	(4) Region FE	(5) Individual characteristics	(6) Marginal costs	(7) Management practices	(8) Rank error	(9) Skewness-adjusted z-score of error
Absolute error z-score	0.011	0.010	0.010	0.011	0.011	0.013	0.016		
	[0.086]	[0.086]	[0.086]	[0.085]	[0.085]	[0.083]	[0.081]		
Don't know	-0.13	-0.16	-0.16	-0.15	-0.16	-0.17*	-0.19*	-0.21**	-0.20*
	[0.11]	[0.11]	[0.11]	[0.11]	[0.10]	[0.10]	[0.10]	[0.10]	[0.10]
Management practices: Monitoring and targeting							0.16	0.16	0.16
							[0.11]	[0.11]	[0.11]
Management practices: Other							0.0095	0.0084	0.010
							[0.11]	[0.11]	[0.11]
Rank error (proportion)								0.00071	
								[0.00075]	
Skewness-adjusted z-score of error									0.041
									[0.048]
Sector fixed effects	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Indicator fixed effects	No	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Region fixed effects	No	No	No	Yes	Yes	Yes	Yes	Yes	Yes
Individual characteristics	No	No	No	No	Yes	Yes	Yes	Yes	Yes
Marginal costs and noise controls	No	No	No	No	No	Yes	Yes	Yes	Yes
Management practies	No	No	No	No	No	No	Yes	Yes	Yes
Adjusted R-squared	0.00073	0.0016	-0.00060	0.041	0.044	0.068	0.075	0.077	0.077
Observations [clusters]	2853 [276]	2853 [276]	2853 [276]	2853 [276]	2853 [276]	2853 [276]	2853 [276]	2853 [276]	2853 [276]

Notes: *** denotes significance at 1%, ** at 5%, and * at 10% level. Clustered standard errors are in parentheses. Standard errors clustered at the district government-sector level. All columns report OLS estimates. The unit of analysis in all columns is the claim of the district government-level official. The sample used for the analysis is the set of respondents at the district government level making claims about their own district governments. The dependent variable is the z-score of the service delivery outcome for the district corresponding to the claim. Sector fixed effects are indicators for whether the claim is within an agriculture, education, health, revenue, or trade organization. Region fixed effects are indicators for the region of the claim. Indicator fixed effects are indicators for the specific service-delivery indicator in the claim and the outcome. Individual characteristics includes: years in position; years in the organization; years in the civil service; an indicator for whether the respondent has an undergraduate degree; and an indicator for whether the respondent has a masters degree; an indicator for whether the respondent is female; and an indicator for whether the respondent is a manager. Marginal costs include the proportion of managers in the organization that state that a management information system is in place; the organization average response to "In what proportion of projects would you say information flows effectively between individuals in the same organization?"; the organization average response to "In what proportion of projects would you say Information flows effectively between organizations?"; the organization average response of managers to "How many personnel do you manage?"; the organization average response of employees to "How many people would you say regularly give you tasks as part of your formal work duties?"; alignment (the z-score of the row mean of the organization-average response to "To what extent do you believe that your organization's mission is aligned to your own mission?" and the organization proportion of selecting "I am contributing to fulfilling that mission on an everyday basis" to "How consistent with your mission are the various tasks and activities assigned to you on a day-to-day basis?"). Noise controls include the time of day of the survey, day-of-survey fixed effects, enumerator fixed effects, an indicator for the enumerator's subjective assessment of the quality of the interview; an indicator for the duration of the interview (decile of duration fixed effects); controls for the average tenure of the managers in the organization and the proportion of female managers in the organization. Management practices incude Management practice: Monitoring and Managament practice: Other. Management practice: Monitoring is the organization average of the monitoring topic, construced from the z-scores of each individual item noted in table A3. Management practice: Other is the organization average of the average across the z-scores of targeting, incentives, roles, flexibility, staffing, and staff involvement, detailed in table A3. Figures are rounded to two significant figures.

Table 38: Comparison of Treated and Untreated Characteristics

Table A11: Balance Table

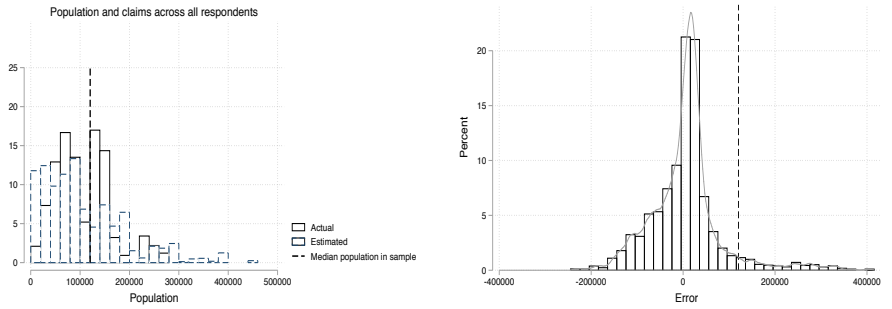
Means and standard deviations

	(1) Control	(2) Treatment	(3) T-test [1]=[2]
Official's Characteristics			
Age	34.12	34.04	0.08
	[8.34]	[8.73]	[0.52]
Years in position	2.92	2.54	0.38**
	[3.03]	[2.34]	[0.16]
Years in organization	7.69	7.34	0.34
	[7.10]	[7.24]	[0.44]
Years in civil service	12.04	12.46	-0.42
	[8.25]	[8.90]	[0.53]
Number of different organizations in service	1.32	1.46	-0.14
	[1.52]	[1.97]	[0.11]
Grade	7.26	6.66	0.60
	[12.59]	[10.22]	[0.69]
Education [undergraduate degree=1]	0.76	0.76	0.00
	[0.43]	[0.43]	[0.03]
Education [masters degree=1]	0.01	0.01	0.00
	[0.12]	[0.11]	[0.01]
Gender [female=1]	0.19	0.17	0.02
	[0.39]	[0.38]	[0.02]
Manager	0.27	0.26	0.01
	[0.44]	[0.44]	[0.03]
Number of bureaucrats	441	679	1120
Organizational Characteristics			
Number of bureaucrats per organization	4.43	4.44	-0.01
	[0.79]	[0.91]	[0.10]
Number of heads per organization	0.95	0.97	-0.02
	[0.21]	[0.16]	[0.02]
Number of managers per organization	0.93	0.89	0.04
	[0.26]	[0.32]	[0.03]
Number of employees per organization	2.54	2.58	-0.03
	[0.73]	[0.78]	[0.09]
Ratio of employees to heads	2.56	2.58	-0.02
	[0.71]	[0.78]	[0.09]
Ratio of employees to managers	2.52	2.61	-0.09
	[0.74]	[0.71]	[0.09]
Span of control (employees per manager)	5.64	6.20	-0.56
	[6.06]	[8.43]	[0.90]
Number of organizations	127	196	323

Notes: Standard deviations are in parentheses. The unit of observation for the first panel is the civil servant and for the second panel the organization. Only employees and managers responded to the information module and hence only these individuals are used for the statistics in the first panel. These statistics are based on the survey sample and not on administrative data. The span of control is taken from the survey question "How many personnel do you manage?"; the total is taken for each organization if there is more than one manager; and the organization-level statistics are reported (each organization has an equal weight in the `By Organization' statistics). Heads refers to heads of organization or deputy heads of organization. `Manager' refers to managers within the organization. Grade is the official civil service grade of the professional civil servant, ranging from 1 to 17 in the sample. The 2013-14 National Civil Service Human Resource Statistics Abstract (Ministry of Civil Service) records the overall female employee percentage to be 35%; this statistic includes frontline staff. Figures are rounded to two decimal places.

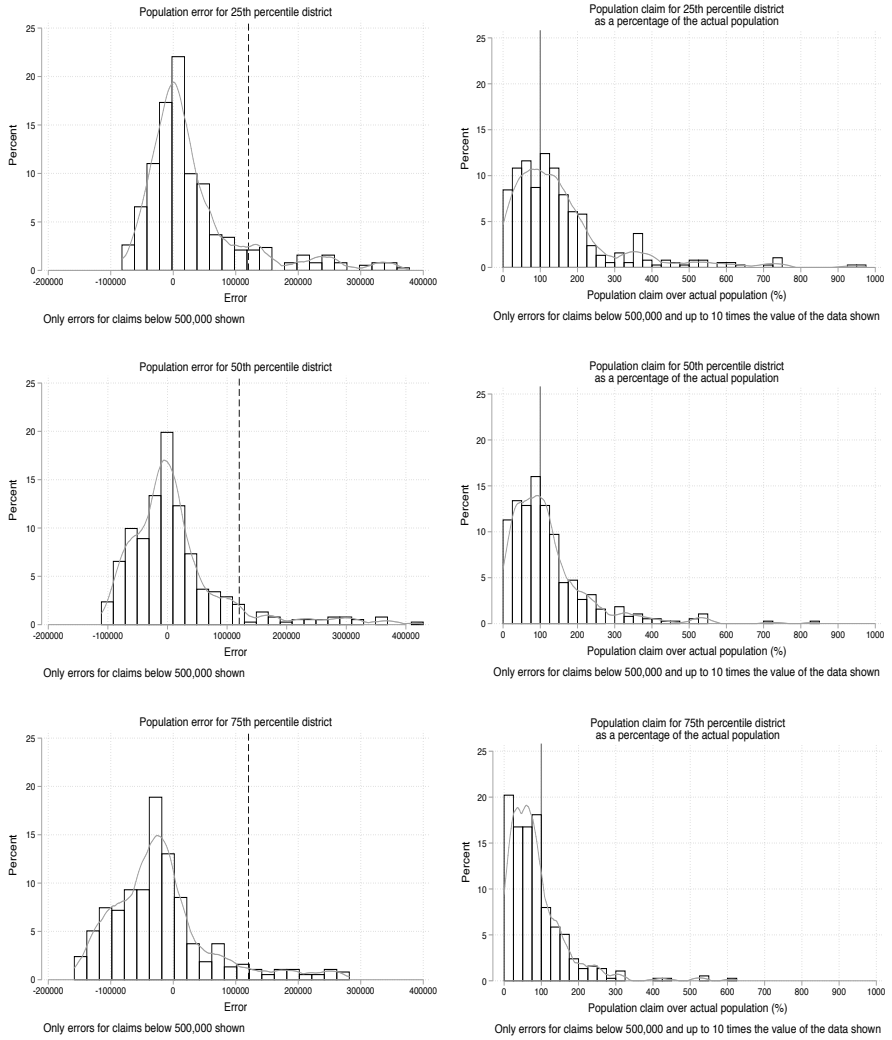
Figure 21: Civil Servants' Estimates on Local Population and Errors

Figure 1A: All Respondents' Population Estimates and Actual Population



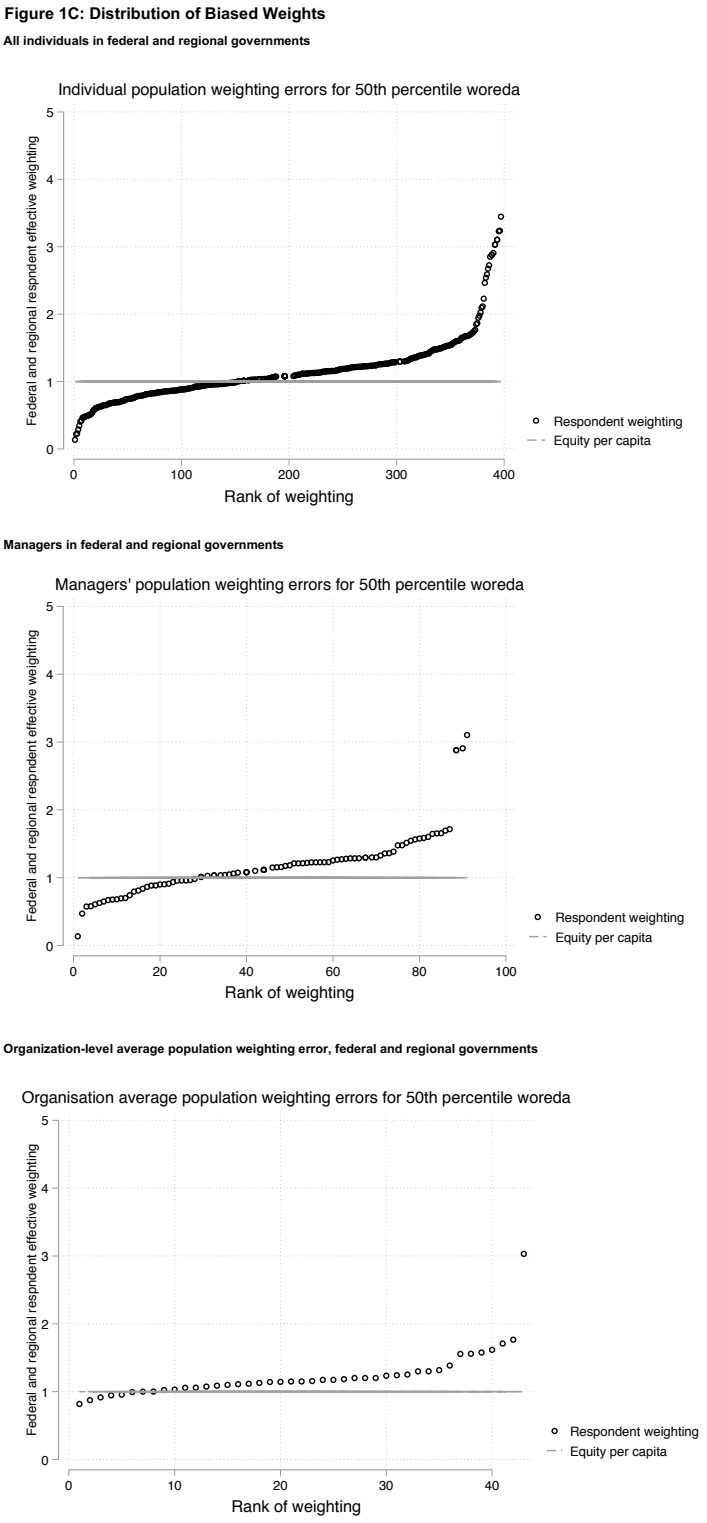
Notes: Figure 1A shows the histograms of actual populations (black) and population claims (blue dashed) for all claims on the left-hand panel; the right-hand panel show histograms of actual errors for all claims. Actual errors are the claim of the respondent minus the actual population value. The black dashed line represents the median actual population in the sample of districts. All claims refers to claims over the 75th percentile, 50th percentile, and 25th percentile district in terms of population within the (chosen) region for the federal and regional respondents; and claims over own districts for district level respondents. The sample includes all employees and managers in the Ethiopian Civil Servants Survey that were asked the information module. In both panels, observations corresponding to populaton claims over 500,000 are not included for presentation purposes.

Figure 1B: Federal and Regional Respondent Population Estimates and Actual Population



Notes: Figure 1B left-hand panels show histograms of actual errors for the 25th percentile district, 50th percentile district, and 75th percentile district. Actual errors are the claim of the respondent minus the actual population value. The black dashed line represents the median actual population in the sample of districts. The sample is all federal and regional employees and managers in the Ethiopian Civil Servants Survey that were asked the information module. The right-hand panel shows histograms of the population claim divided by the actual population of the district multiplied by 100 to convey the size of the population claim relative to the data in percentage terms. The solid gray line at 100 represents an accurate claim. In both panels, observations corresponding to populaton claims over 500,000 are not included for presentation purposes. In the right-hand panel, observations where the claim is more than 10 times (1000% of) the actual data are not shown.

Figure 22: Civil Servant Errors and Hypothetical Effects on Budget Allocations



Notes: The figure shows the organization-average estimated population weights for the median district in the region in terms of population. The top panel shows the observations for all employees and managers in the sample; the middle panel shows the observations for all managers in the sample; the bottom panel shows the organization-level averages of all employees and managers in the sample. The weights are calcuated based on respondent estimates of the population across the 25th, 50th, and 75th percentile district in the region. The black line represents equity per capita -- the correct weights if the actual data were used; the black circles represent the weight for the median disrict based on the region and federal civil servant respondents' estimates of the populations of the 25th, 50th, and 75th percentile districts. The organization average weight is then used for the purpose of this figure. The sample for the top panel is all federal and regional employees and managers that responded to the information module in the Ethiopian Civil Servants Survey; the sample for the middle panel is all federal and regional managers that responded to the information module in the Ethiopian Civil Servants Survey; the sample for the bottom panel is all federal and regional employees and managers that responded to the information module in the Ethiopian Civil Servants Survey.

Figure 23: Decentralisation and Information

Figure 2: Distributions of Errors Across Tiers of Government

Figure 2A: CDF of Errors Using Z-Scores

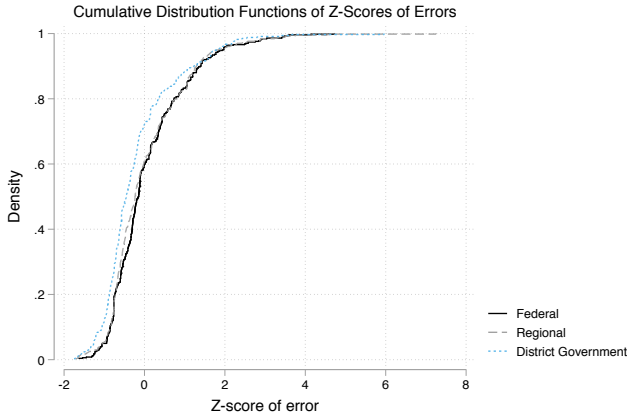


Figure 2B: CDF of Errors Using Skewness Adjusted Z-Scores

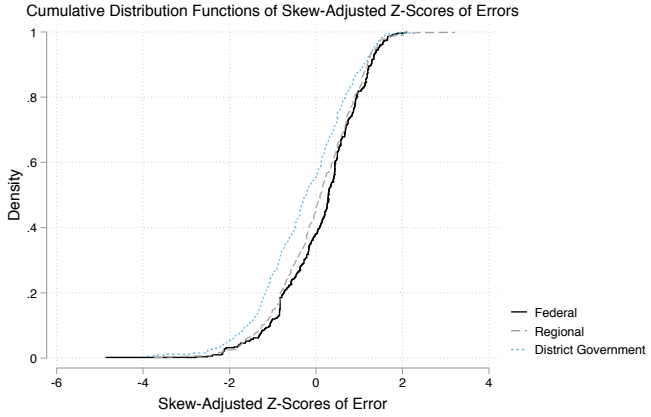
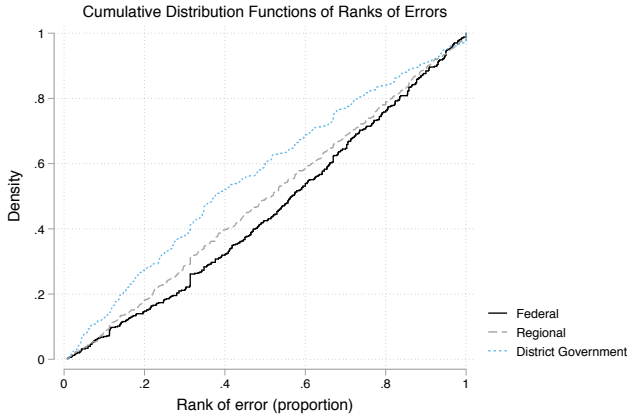


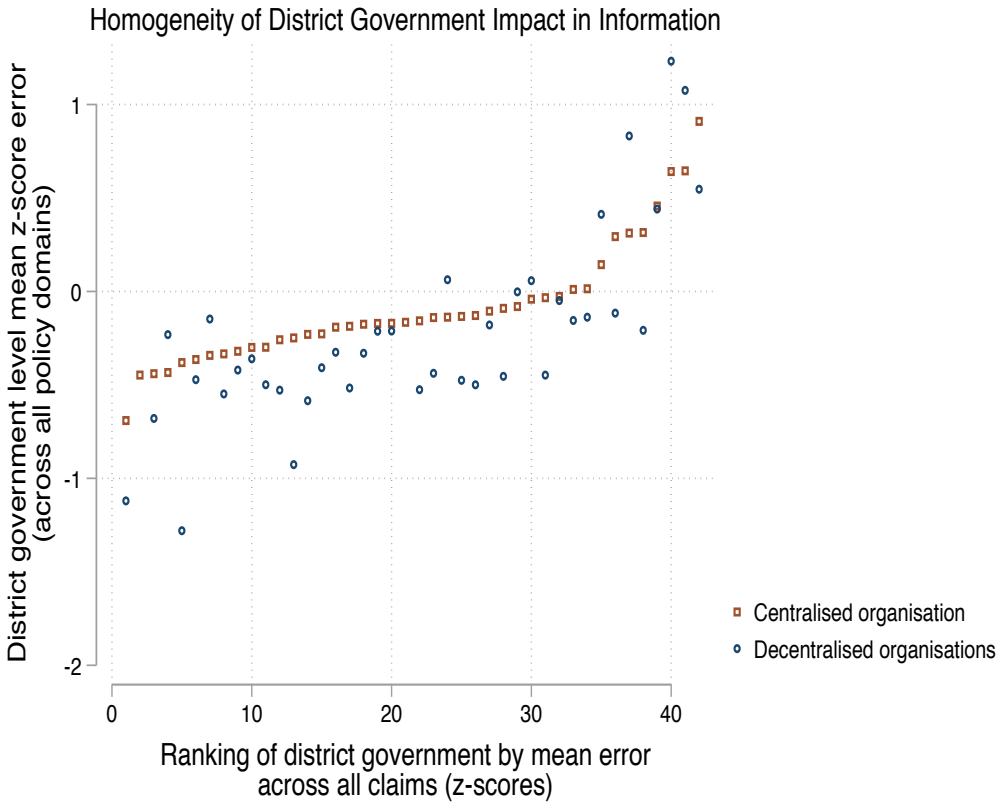
Figure 2C: CDF of Errors Using Ranks



Notes: The figure shows the cumulative distribution functions of the federal, regional, and district government level respondent's errors across all policy domains. The top panel shows the errors in terms of z-score within the indicator. The middle panel shows the error in terms of skewness-adjusted z-score within the indicator; the skewness-adjusted z-score is the standard z-score in cases of variables with a 0 to 1 support and for variables that are not significantly skewed according to the test from D'Agostino, Balanger, and D'Agostino (1990); for those variables that are not 0 to 1 in support and are significantly skewed, the log of the variable is taken and the z-score of this is the skewness-adjusted z-score. The bottom panel shows the errors in terms of rank within the indicator as a proportion of the maximum rank. The sample used in the analysis is the set of respondents making claims about the same district. The black solid line refers to federal claims; the gray dashed line to regional claims; and the blue dotted line to district government claims.

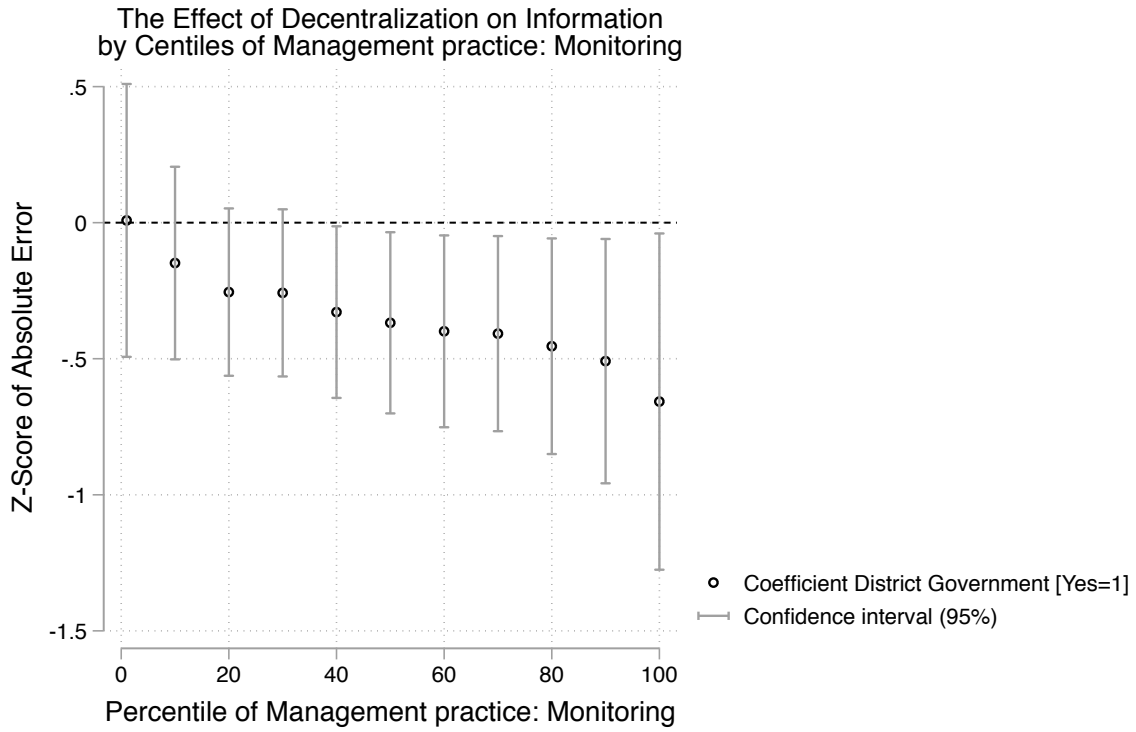
Figure 24: Heterogeneous Effects of Decentralisation on Information Across Jurisdictions

Figure 3: Assessing Homogeneity of District Impact



Notes: The figure shows the average claims for centralized (federal and regional) organizations and decentralized (district government) organizations, averaged by the district government that they are claiming over. The red squares refer to centralized organizations and the blue circles to decentralized organizations. The sample used for the analysis is the set of respondents that make claims about the same districts. The error is in terms of z-score within the indicator.

Figure 4: Incentives, Decentralization and Information

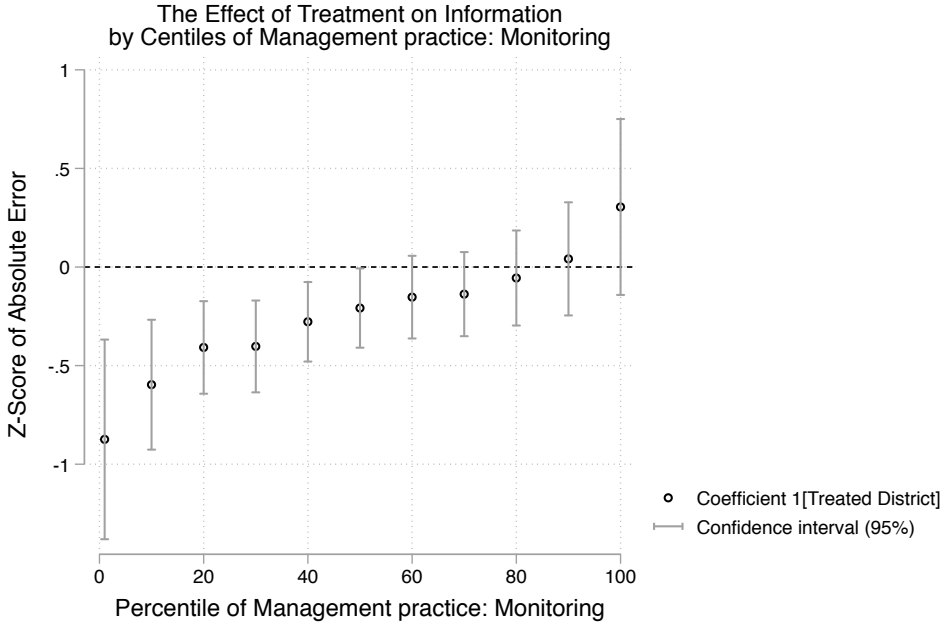


Notes: The figure shows the OLS coefficient on District Government [Yes=1] for each decile of Management practice: Monitoring, in a regression also conditional on Management practice: Other; Individual controls; Sector fixed effects; Indicator fixed effects, Percentile fixed effects, Marginal cost controls and Noise controls. 95% confidence intervals are show in the gray lines, with standard errors clustered at the district government-sector interacted level.

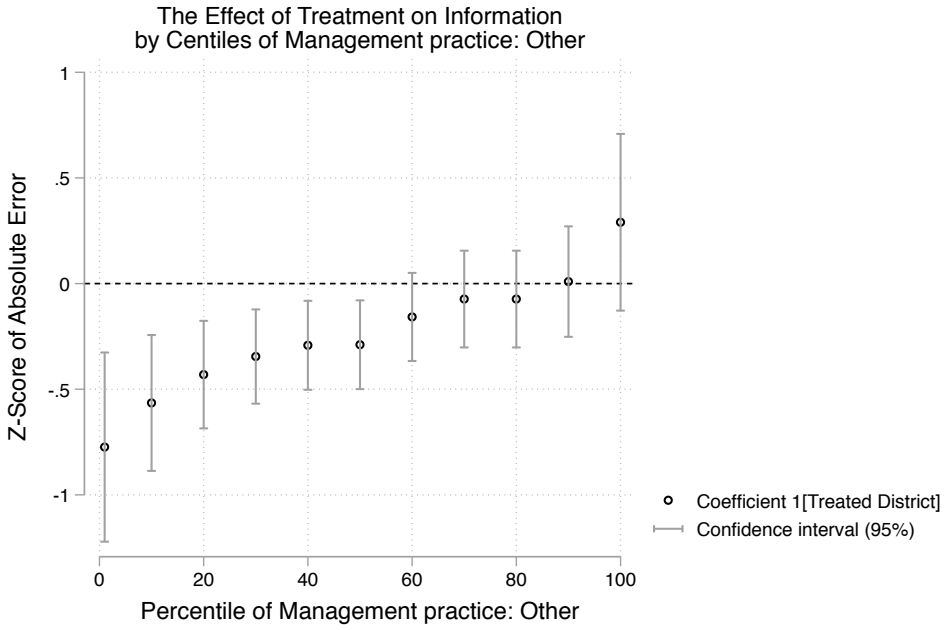
Figure 26: Heterogeneous Effects of the Treatment by Management Practices

Figure 5: Heterogeneous Treatment Effects by Management practices

5A: Treatment and Management practice: Monitoring



5B: Treatment and Management practice: Other



Notes: The figure shows the effect of the treatment on the z-score of errors acorss all policy domains by different percentles of Management practice: Monitoring in the top figure; and by different percentiles of Management practice: Other in the bottom figure. The sample used for the analysis is the set of respondents that make claims about the same districts. Management practices and errors are in z-scores. Treated district is a binary indicator equal to one if the district received an information package Regressions are conditional on Management practices; Individual controls; Sector fixed effects; Indicator fixed effects, Percentile fixed effects, Marginal cost controls and noise controls, Region fixed effects; and an indicator for whether the claim is from a District Government. 95% confidence intervals are shown in the gray lines, with standard errors clustered at the district government-sector level. Management practice: Monitoring is the organization average of the monitoring topic, construced from the z-scores of each individual item noted in table A3. Management practice: Other is the organization average of the average across the z-scores of targeting, incentives, roles, flexibility, staffing, and staff involvement, detailed in table A3.

Figure 27: A Map of Sampled District Governments

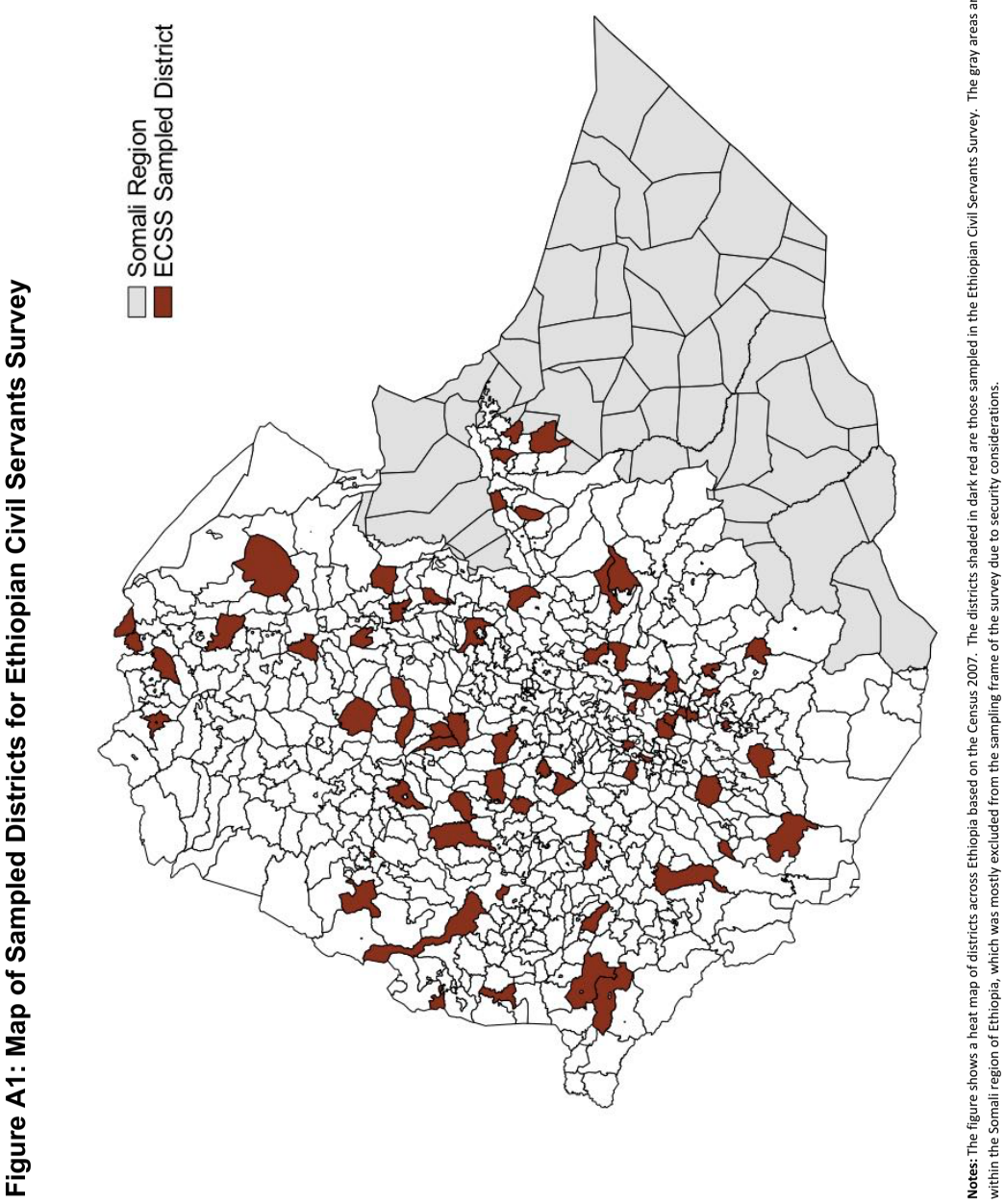
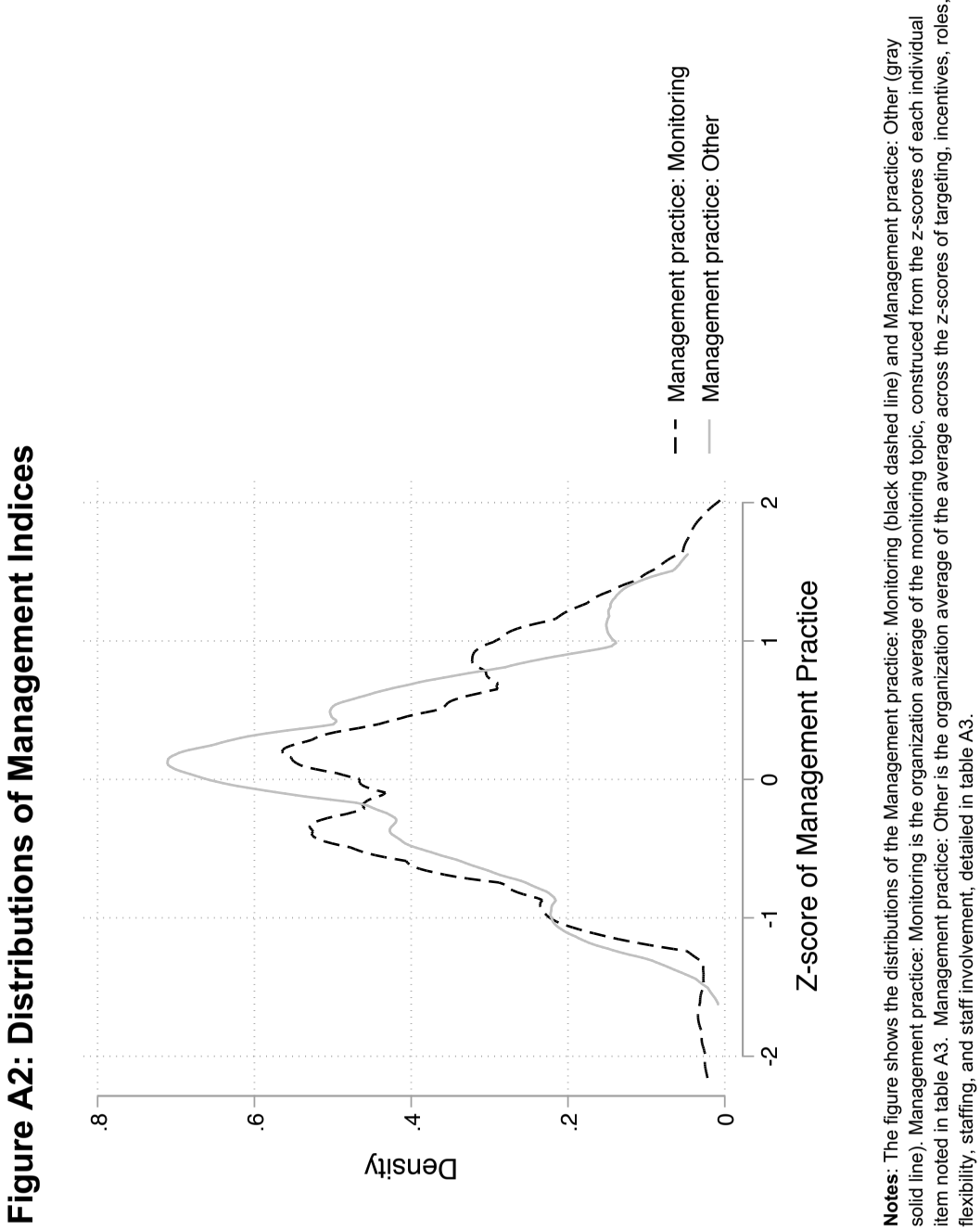


Figure 28: Distributions of Management Practice Indices Across Organisations



Notes: The figure shows the distributions of the Management practice: Monitoring (black dashed line) and Management practice: Other (gray solid line). Management practice: Monitoring is the organization average of the monitoring topic, construced from the z-scores of each individual item noted in table A3. Management practice: Other is the organization average of the average across the z-scores of targeting, incentives, roles, flexibility, staffing, and staff involvement, detailed in table A3.

4 Does Providing Information to Civil Servants Improve Service Delivery?

4.1 Introduction

The question of how to improve the ability of public organisations to transform scarce resources into vital public goods has been long-studied in the social sciences [Rose-Ackerman, 1986; Pepinsky et al., 2017]. The decisions of bureaucrats, who allocate scarce time, money and effort across different policy dimensions, are an important part of this process [Dixit, 2002*b*; Meyers and Vorsanger, 2007].

Deciding optimal allocations across a multitude of different dimensions requires detailed knowledge of the production function of public goods; however, there is strong evidence that civil servants have poor information even on existing levels of service delivery indicators [Rogger and Somani, 2018]. This could be the result of limited access to information technology [Rogger, 2017], limited access to reliable data [Boräng et al., 2018], or as argued by Duflo [2017], a lack of "time or inclination." By analysing the introduction of an e-governance intervention that provides public officials with timely information on the conditions in their local jurisdiction, this paper provides empirical evidence that bureaucrats do respond to new information and that this feeds through to public-service delivery.

This paper takes advantage of a new district-level panel dataset on the internal functions of local governments in Ethiopia. The data captures the phased introduction of an intervention that allows civil servants to access an electronic database on service delivery in their jurisdiction. The intervention is rolled out such that those districts closest to the regional capitals receive the programme sooner, which ensures that, conditional on fixed effects (including distance from the regional capital), the effect of the programme on service-delivery outcomes can be identified. In a context where information on service delivery is severely limited (for example see Rogger and Somani [2018]), the intervention results in a drastic improvement in the access to such information, allowing bureaucrats to adjust the allocation of money, time and effort across different activities. These decisions of local education-sector bureaucrats feed directly into education policy planning and implementation decisions, such as budget allocations, the hiring and firing of teachers, organisational management, and implementation effort. I find that the intervention leads to significant improvements in enrolment (access) but significant short-run deteriorations in schooling conditions, consistent with civil servants reallocating investments across activities in response to the new information.

I put forward a simple theoretical model that shows that bureaucrats will respond

to the intervention if it provides new information regarding the expected returns of individual activities. Before the intervention, bureaucrats have beliefs over the returns to individual inputs in the production of public services. After the intervention, bureaucrats update their beliefs based on the new and improved information. Risk-averse bureaucrats will optimally allocate more resources to activities with larger average returns or lower variability in returns, relative to prior beliefs. If the budget remains fixed, certain inputs will increase, while others will (optimally) decrease.

The empirical analysis provides evidence of overall positive effects of the intervention on education *access*. The results suggests that, on average, access to the technology within district education offices leads to an increase of 2.7 percentage points in the enrolment of 7-14 year-old children in primary school, which is 3.2% relative to the baseline mean, or 18% of the remaining gap towards universal primary education, or an additional 167,262 children in primary school.

However, consistent with the theoretical framework, I find evidence that indicators relating to schooling conditions significantly worsen, providing evidence of multitasking concerns. The empirical results demonstrate significantly worse pupil-teacher ratios and pupil-class ratios as a result of the intervention, as enrolments expand. The worse schooling conditions occur as a result of reductions in teacher hiring and classroom construction, suggesting that public officials trade off investments aimed at increasing schooling access at the expense of investments in schooling conditions. It takes 2 years after the programme for schooling conditions to return to pre-intervention levels, as more teachers are hired and classrooms constructed. While these negative effects are a concern, the new allocations may be pareto efficient, as public officials re-optimise their allocation decisions in response to new, improved information on the returns to investments in different activities.

The analysis of the heterogeneous effects of the intervention shows that those districts facing more severe information constraints over the returns to individual activities prior to the intervention, as proxied by changes in enrolment numbers during the pre-intervention period, display larger responses to the intervention. These results provide further evidence that information is indeed a constraint to public-service provision.

I investigate the possible channels generating the results. I rule out an interaction between information and formal performance incentives; and rule out changes in the budget-bargaining mechanism between regional and district governments, and therefore changes in financial resources. The empirical results are consistent with the conceptual framework of civil servants optimally reallocating fixed resources in response to new information about the returns to investments. Using supplementary data from the Ethiopian Civil Servants Survey 2016, I provide descriptive evidence that the service delivery information is used in the *subjective* performance evaluations of the civil

servants. This provides evidence that the intervention affects service delivery through non-monetary incentives and career concerns.

Information on the implementation costs shows that the programme costs around 35 USD per additional pupil enrolled as a result of the programme, without taking into account the costs of hiring additional teachers or constructing additional classrooms to maintain the quality of schooling conditions. Taking into account these costs implies a cost of 54 USD per additional pupil enrolled. This cost requires low lifetime returns to education per pupil to pass any reasonable cost-benefit evaluation, but is still larger than the current expenditure per pupil (30 USD).

A natural concern for identification is that the common-trends assumption fails: districts that received the intervention earlier display differing trends in the outcome variables. I use pre-intervention data to show that there is no evidence of such concerns in the data. An alternative concern is that the receipt of the intervention coincides with a time-varying omitted variable that affects the provision of public services. Using information on service-delivery indicators in the health and water sectors, I show that the intervention has no effect on non-education service delivery outcomes, providing evidence against such concerns. While the dataset that I have does not capture the specific investment decisions of individual-level civil servants, I have district-level data on education service inputs, such as teachers and classrooms, which provides information on allocation decisions at the district level.

This paper contributes to the literature on the productivity of public officials [Rose-Ackerman, 1986; Finan et al., 2017; Pepinsky et al., 2017]. While significant strides have been made in understanding the role of management in public service delivery [Rasul and Rogger, 2016], the effects of monetary incentives on the selection of bureaucrats [Dal Bó et al., 2013; Ashraf et al., 2016], and the effects of performance incentives on frontline service providers [Glewwe et al., 2010; Basinga et al., 2011; Muralidharan and Sundararaman, 2011; Olken et al., 2014; Khan et al., 2015], there remain gaps in understanding alternative mechanisms to improve the performance of employed bureaucrats, particularly those working in the middle tiers of government. This paper provides empirical evidence that bureaucrats respond to better information regarding their operating environment by reallocating investments across activities. This evidence demonstrates that bureaucrats re-optimise in response to new information and that multitasking concerns are prevalent in the provision of public services with fixed budget constraints.

Governments have been spending large amounts on e-governance investments in recent years (for example over $600 billion by the Federal Government of the United States in the past decade alone) and this trend is increasing; yet only a very small proportion of these projects are recorded as successes [World Bank, 2016*c*]. Existing research

highlights the effects of e-governance on the attendance of public officials [Dhaliwal and Hanna, 2017], the extent of corruption and the efficiency of public programs [Miller and Tucker, 2011; Barnwal, 2014; Muralidharan et al., 2016; Banerjee et al., 2015], electoral fraud [Callen and Long, 2014], the quality of procurement contracts [Lewis-Faupel et al., 2016], the experience of service users [Aker et al., 2016], and voter representation [Fujiwara, 2015]. While existing research suggests largely positive effects from e-governance on a multitude of dimensions, the literature to date is scarce and concentrated in a handful of countries. This paper contributes to this research agenda and uses detailed cost data to conduct a cost-benefit analysis of the intervention.

The paper is organised into the following sections: 5.2 describes the context of governance in Ethiopia, the education sector, the technology, and the dataset; 4.3 outlines the theoretical framework used to guide the empirical analysis and interpret the results; 4.4 outlines the identification strategy and the results; 4.5 investigates potential mechanisms; 4.6 presents robustness checks; 4.7 concludes.

4.2 Context, programme details and the data

4.2.1 The government structure in Ethiopia

Governance in Ethiopia is centred around three major tiers of government: federal, regional, and district (woreda). Some of the regions have zones, which act as intermediaries between the regional- and district-level governments; kebeles are also considered part of the government structure and can be thought of as neighbourhood collectives that interact with both the district governments and the communities [Federal Democratic Republic of Ethiopia, 2012]. Since a major drive to decentralise budgets, decision making, and revenue collection through the *District Level Decentralization Program*, the majority of the planning and design of non-federal economic policies is conducted at the district level, with districts considered the socio-economic centres of policy [Ministry of Finance and Economic Development, Ethiopia, 2002; Gebre-Egziabher, 2007; Khan et al., 2014].

Regions are transferred close to half of all public resources and, of this, districts manage around 80% [World Bank, 2010*a*, 2015*d*]. Almost all of the non-federal expenditure, therefore, flows through the districts, which have *de jure* autonomy in its allocation. This makes the district-level government the major tier when it comes to non-federal policy.

4.2.2 The education sector in Ethiopia

The education sector in Ethiopia is the largest sector-recipient of public money, receiving 22% of annual public expenditure (around 5% of GDP) in 2010 [World Bank, 2015*b*]. More recent expenditure data from the Ethiopian Treasury shows that the education sector accounted for 33% of all non-federal public expenditure in 2013/14 – the largest single item – and 38% at just the district level – again, the largest single item. This means that one-third of all non-federal expenditure is dedicated to education. The sector is a major focus of the current government strategy and receives much attention from top policymakers in the country [Ministry of Finance and Economic Development, Ethiopia, 2010]. Over the recent past, the government has achieved significant improvements in the education sector in Ethiopia, particularly towards attaining the ambitious target of education for all, with a current net-enrolment rate of around 85% [United Nations Educational, Scientific and Cultural Organization, 2012; Ministry of Education, Ethiopia, 2015]; although there remain some concerns over the quality of schooling as the number of enrollees increases [Engel, 2010; World Bank, 2013*b*; Zike and Ayele, 2015].[55]

4.2.3 District level government education policy

The districts are autonomous units in the Ethiopian government, able to determine local policy subject to the preferences of their constituents: "While Ethiopia is a federal state, each region and district has autonomy which is ensured by the constitution. Decentralization brings service providers under the control of local governments and their constituents" [p.4, UNESCO, 2015]. The district is responsible for the majority of non-federal education policy implementation: "The latest decentralisation reform, which started over a decade ago, has transferred important responsibilities for general education to district offices. These offices now exercise their responsibilities, with support from regional offices, within an overall framework developed at the federal level" [p.119, Federal Ministry of Education, 2015].

The district is responsible for almost all major implementation of education policy, such as budget allocation decisions; hiring and firing of education personnel; physical infrastructure of the schools; monitoring local education progress; community engagement; transport provision; school-feeding programmes; and other capital-investment decisions.[56]

[55]For summary performance statistics published by the Ministry of Education on the state of national-level and regional-level education indicators, see Ministry of Education, Ethiopia [2015]

[56]Individual school institutions are responsible for all decisions regarding school management and coordinate with the district office regarding resource requirements.

4.2.4 The Education Management Information System (EMIS)

A significant part of the government's strategy to improve the quality of education provision has been through improving the availability and quality of information to policymakers and civil servants in the education sector. This is primarily conducted through the Education Management Information System (EMIS) programme [World Bank, 2013*a*]. The Ministry of Education identified the need for better information over the state of provision, particularly regarding certain service delivery indicators, which are considered key measures of provision quality – such as enrolments, pupil-teacher ratios, and pupil-class ratios – in order to aid policymaking and planning [Ministry of Education, Ethiopia, 2010].[57]

As set out by UNESCO, the objectives of an EMIS include: improving capacities in data processing, storage and analysis; facilitating and promoting the use of relevant information, at all levels, for more effective education planning and implementation; and providing information for policy dialogue [Villanueva, 2003]. The national education strategy states that "EMIS will be the primary source for monitoring data for each sub-sector during ESDP V" [p.121, Federal Ministry of Education, 2015]. The national education strategy outlines the importance of the district-level EMIS programme in terms of monitoring schooling progress and acting upon this information through policy-planning responses: "Woreda [District] EMIS offices will analyse local data and provide school leaders with up-to-date performance information from EMIS and inspection findings, relative to schools in their woreda and to national standards ... By producing report cards, woreda officials will be better placed to identify strengths and weaknesses and to use this for targeted and timely responses" [pp.48-49, Federal Ministry of Education, 2015]. The role of the EMIS, as the primary source of monitoring data, is therefore to track education progress in the district and inform policy.

Prior to such a system being in place, this information would only be available to civil servants through a paper-based system, which was often not utilised due to the time and processing costs of doing so: "[in] the absence of a computerised database, woreda education plans are based on a relatively restricted range of disaggregated data because it is infeasible to analyse a mass of unrelated manual data" [p.78, Ministry of Education, Ethiopia, 2010].

[57] The pupil-class ratio is referred to pupil-section ratio in the documents of the Ministry of Education, Ethiopia. Classrooms in the Ethiopian system refer to the physical structure of the classroom itself. Sections refer to the size of individual classes. Classrooms can be used twice in a day by different school cycles (e.g. grades 1-4 in the morning and grades 5-8 in the afternoon) due to double school shifts. A school, then, can have 10 classrooms and 20 sections if it uses a double-shift structure.

4.2.4.1 The data

The dataset used in this analysis is unique. The dataset is collected from within government in Ethiopia and provides a rare insight into internal government functions. The education data was collected directly from the official database of the Ministry of Education in Ethiopia, extracted onsite. This database contains information on 707 districts outside of Addis Ababa.[58] Information on the EMIS database training programmes was gathered by the Ministry of Education from the regions and districts for the purpose of this study.[59] All population and demographic data are taken from the 2007 Ethiopian census.[60] District-level budget information is collected from the World Bank Open Budget Portal initiative.[61] The Woreda and City Benchmarking Survey covers a subset of districts and captures information on district institutional characteristics. District-level demographic information is taken from the census of 2007 and the World Bank Ethiopia's Poverty and Social Impact Analysis Database [Khan et al., 2014].[62] A description and the coverage of the main variables used in this paper are presented in table 47. The descriptives presented in section 4.5.3 are drawn from the Ethiopia Civil Servants Survey of 2016, which is used as a standalone dataset in this analysis, not merged with any of the above.

Data on the costs of the programme are extracted from annual work plans from the General Education Quality Improvement Programme of the Ministry of Education.[63]

Unfortunately, not all districts are observed in all years of data for all indicators. This analysis focuses only on those 503 (71%) out of the 707 districts that are observed in all years of the data for all indicators. The most recent pre-intervention period (2008/09) characteristics of the included and excluded districts are shown in table 48. When comparing the districts for which we have data for all years to the total set of districts in 2008/09, we see that the districts used in this analysis are representative of the population of districts as a whole, due to the large coverage of the population of districts in the sample database.

[58]The census of 2007 contains 726 districts, excluding the special administration areas and Addis Ababa

[59]I would like to thank Bereket Kelemu, Debebe Wordofa, Gabi Elte, and Jack Rossiter in the Ministry of Education for all their help in accessing and extracting the data.

[60]http://www.csa.gov.et/census-report/complete-report/census-2007 [Accessed 31 July 2018]

[61]World Bank Open Budget Portal initiative: http://wbi.worldbank.org/boost/boost-initiative [Accessed 31 July 2018]

[62]The data was cleaned, merged, and put into the dataset used for the current analysis, with the merging done by region-district identifiers.

[63]http://www.worldbank.org/projects/P129828/ethiopia-general-education-quality-improvement-project-ii?lang=en [Accessed 31 July 2018]

4.2.5 The programme

All district offices had the technical infrastructure for EMIS procured by 2008/09, which places all district offices at the same level in terms of potential access to the new technology prior to the start of any training. Once the infrastructure is rolled out to the district offices, there is still limited access to the EMIS without the appropriate database training, which the evaluated programme provides. This is because of the need to install, set up, and train on the use of Microsoft Access to upload and manage the education data within a database, and also on the use of a specific education database package used by policymakers to extract specific education information from the database (StatEduc2).[64] The evaluated programme therefore 'unlocks' the use of the EMIS infrastructure by local-level civil servants. The programme is focused specifically on how to organise and use this data; for example, a document from the Ethiopian Civil Service University outlines one particular training for EMIS at the university:[65] "exercise group members tried ways of organizing, classifying, handling data to generate a user friendly and informative report"..."[The director of Ethiopian Civil Service University's Center for Academic and Professional Development] expressed hope the training had enabled trainees [to] internalize the importance of properly handling data" [Ethiopian Civil Service University, 2012]. The programme, therefore, is rather narrow in its purpose and this should limit any concerns regarding spillover effects from the database training.

Table 39 shows the rollout of the programme. The selection of districts to receive the intervention is based on district-specific costs of access for the regionally-provided trainers, such that those districts further away from the regional capital receive the programme at a later date. The time-varying introduction of the programme allows the identification of the average treatment effect on the treated (4.4.1) and also an analysis of lagged effects (4.4.4)

4.2.6 Channels of effects

Figure 29 illustrates the channels of effects that are identified by the empirical strategy. The infrastructure is procured and rolled out nationally by 2008/09 across all districts; the database training on how to access and use the EMIS is rolled out from 2009/10 to 2013/14, based on district distance from the regional capital. The channel of effects that the empirical strategy aims to identify is therefore: the database training 'unlocks' the infrastructure to access and make use of the EMIS data; the district-level civil

[64] For example, the pupil-teacher ratio for the district, or for each school in the district

[65] Though not for district-level civil servants specifically, the purpose of the training is the same: to train on the use and application of EMIS

servants have much easier access to information regarding the state of service delivery in the district and therefore know the conditions in and productivity of the district with greater accuracy; this informs the civil servants of the performance of the district in terms of education provision, using the main indicators on which education progress is tracked; the civil servants then adapt policy planning decisions at the district level, as described above; this then feeds through to education provision outcomes.

4.2.6.1 Measures of 'performance'

The purpose of this paper is to evaluate the effectiveness of this programme on the specific service delivery indicators that the EMIS **aims** to impact. Although the outcomes used to evaluate these programmes do not focus on the ultimate goal of education – learning outcomes – these service delivery indicators are those that the government uses to evaluate its own performance in terms of providing quality education, and therefore the areas in which there exist pressure from top-tier policymakers and development partners [Ministry of Education, Ethiopia, 2015; World Bank, 2013*a*; United Nations Educational, Scientific and Cultural Organization, 2012].

4.3 Conceptual framework

4.3.1 Outline

The conceptual framework presented here is a parsimonious model that relates directly to the expected effects of the intervention in question and is by no means a complete model of the production of public goods with incomplete information.

I assume that the intervention functions through allocation decisions of managers of public organisations, who I allow to be risk averse and altruistic.

The final public good (e.g. 'quality education provision') is a function of two (without loss of generality) activities a_1 and a_2. These activities could be investments in the provision of classrooms, investments in the provision of qualified teachers, investments in the provision of school facilities, or investments in other activities to reduce the cost of attending education, for example the provision of free school meals or school buses (see section 4.2.3).[66] I do not observe the size of the investments or the time allocation of civil servants, but I do observe the final provision of classrooms and teachers, as well as enrolments.

The managers decide on which activities to focus resources, based on their beliefs over the returns to each activity. I assume that the managers have tacit beliefs over the relative returns to investing more time, money, and effort in a_1 relative to a_2,

[66] Investments can be in time, effort, or monetary investments.

represented by θ_1 and θ_2 respectively. These beliefs are determined by the information available to the managers.[67] The production function from activities to the public good is then:

$$Q = \theta_1 g_1(a_1) + \theta_2 g_2(a_2) \qquad g_i' > 0 \;\; g_{ii}'' < 0 \;\; i = \{1, 2\}$$

Where $g_i(a_i)$ is the continuous mapping from the activity i to the public good Q, without accounting for district-level productivity θ_i for that activity.[68] θ_i represents local conditions which affect the overall efficiency with which activity a_i can be transformed into public goods. These can be, for example, the local availability of skilled graduates, the local infrastructure for children to reach local schools, the local availability of materials to build classrooms and schools, local demand for schooling etc.[69].

I assume that the manager knows a_i and $g_i(.)$, but is uncertain about the relative returns to each activity, θ_i, $i = \{1, 2\}$. Managers have beliefs over, θ_i, represented $\hat{\theta}_i$, distributed independently as follows:

$$\hat{\theta}_i \sim \;\; i.i.d \; N(\mu_i, \sigma_i)$$

It is the distribution of beliefs, $\hat{\theta}_i$, that the EMIS impacts. The intervention provides new, more accurate information regarding the output Q. In combination with knowledge over $g_i(a_i)$, $i = \{1, 2\}$, the manager is able to learn the true θ_i with greater accuracy.

I allow the contracting environment of the manager to closely follow the case of the civil service in Ethiopia, assuming fixed wages (w). The manager weighs the level of public goods Q in her utility function by $\eta > 0$, allowing for altruism, career concerns, and/or performance incentives in the preferences or contracts of managers.

Managers have risk-averse preferences represented by a constant-absolute-risk-aversion utility function $u(x) = -\exp(-rx)$, where r represents the risk-aversion of the managers.

I assume that the manager is provided a fixed budget, B, which she is allowed to allocate across activities freely. The resources allocated to the two activities must not be greater than the amount of resources available: $a_1 + a_2 \leq B$, and the total resources (B) are fixed over time – I provide empirical evidence for this assertion in section 4.5.2. In optimum, in the absence of any other investment products, the manager invests all

[67] The information was largely disaggregated masses of paper prior to the EMIS intervention – see section 4.2.4

[68] Note that the production function for public goods assumes additive separability; this strong assumption is not necessary for the broad qualitative implications of the effects of the intervention (through the comparative statics) to hold; however, the inclusion of complementarity in the production function creates multiple equilibria which complicate the comparative static solutions.

[69] I do not observe these conditions.

resources in the two activities: $a_1 + a_2 = B$; or $a_2 = B - a_1$.

The optimisation problem of the manager under these assumptions is:

$$Max_{\{a_1,a_2\}} \; E\Big(- \exp(-r\eta\theta_1 g_1(a_1) - r\eta\theta_2 g_2(a_2) - rw)\Big)$$

Which, using the results of the log-normal distribution, becomes:

$$Max_{\{a_1,a_2\}} \; - \exp(-r\eta\mu_1 g_1(a_1) - r\eta\mu_2 g_2(a_2) - rw + \frac{r^2}{2}\eta^2\sigma_1^2 g_1^2(a_1) + \frac{r^2}{2}\eta^2\sigma_2^2 g_2^2(a_2))$$

Using $a_1 = B - a_2$, the problem simplifies to:

$$Max_{\{a_1\}} \; - \exp(-r\eta\mu_1 g_1(a_1) - r\eta\mu_2 g_2(B-a_1) - rw + \frac{r^2}{2}\eta^2\sigma_1^2 g_1^2(a_1) + \frac{r^2}{2}\eta^2\sigma_2^2 g_2^2(B-a_1))$$

Then using the fact that maximising $-\exp(x)$, with respect to x, is equivalent to minimising $\exp(x)$, and the fact that minimising $-\exp(x)$ is equivalent to maximising $\exp(-x)$, which is equivalent to maximising the certainty equivalence:

$$\{a_1^*\} = arg\, max \left[r\eta\mu_1 g_1(a_1) + r\eta\mu_2 g_2(B-a_1) + rw - \frac{r^2}{2}\eta^2\sigma_1^2 g_1^2(a_1) - \frac{r^2}{2}\eta^2\sigma_2^2 g_2^2(B-a_1)\right]$$

Which, assuming an internal solution with positive a_1, yields the first-order condition:

$$\mu_1 g_1'(a_1) - \mu_2 g_2'(B - a_1) - r\eta[\sigma_1^2 g_1(a_1) - \sigma_2^2 g_2(B - a_1)] = 0 \tag{26}$$

The second-order condition is satisfied under the assumptions of the shape of g, using $T - a_1 = a_2$:

$$\mu_1 g_{11}''(a_1) + \mu_2 g_{22}''(a_2) - r\eta[\sigma_1^2 g_1'(a_1) - \sigma_2^2 g_2'(a_2)] < 0 \tag{27}$$

The intervention of the introduction of the management-information system provides new, up-to-date information on the current level of outputs Q. In combination with knowledge over $g_1(a_1)$ and $g_2(a_2)$, the manager can update her beliefs over the distribution of θ_1 and θ_2. The updated beliefs follow the distribution:

$$\theta_i' \sim i.i.d. N(\mu_i', \sigma_i')$$

Without loss of generality, I provide the comparative statics of the optimal response

of a_1 as the manager updates her knowledge of θ_1 and θ_2.[70] The corresponding response to a_2 is determined by $a_2 = B - a_1$.

The following comparative statics demonstrate the optimal response of the manager's decision of a_1 as μ_1, μ_2, σ_1, and σ_2 increase:

$$\frac{\partial a_1}{\partial \mu_1} = -\frac{g_1'(a_1)}{\mu_1 g_{11}''(a_1) + \mu_2 g_{22}''(B-a_1) - r\eta\sigma_1^2 g_1'(a_1) - r\eta g_2'(B-a_1)} > 0$$

$$\frac{\partial a_1}{\partial \mu_2} = -\frac{g_2'(B-a_1)}{\mu_1 g_{11}''(a_1) + \mu_2 g_{22}''(B-a_1) - r\eta\sigma_1^2 g_1'(a_1) - r\eta g_2'(B-a_1)} < 0$$

$$\frac{\partial a_1}{\partial \sigma_1^2} = -\frac{-r\eta g_1(a_1)}{\mu_1 g_{11}''(a_1) + \mu_2 g_{22}''(B-a_1) - r\eta\sigma_1^2 g_1'(a_1) - r\eta g_2'(B-a_1)} < 0$$

$$\frac{\partial a_1}{\partial \sigma_2^2} = -\frac{r\eta g_2(B-a_1)}{\mu_1 g_{11}''(a_1) + \mu_2 g_{22}''(B-a_1) - r\eta\sigma_1^2 g_1'(a_1) - r\eta g_2'(B-a_1)} > 0$$

4.3.2 Implications

Let a_1' be the post-intervention allocation decisions of the manager. The key implications from the conceptual framework are summarised:

1. If, ceterus paribus, the intervention provides information that the average returns to a_1 are higher than her prior belief, $\mu_1' > \mu_1$, then $a_1' > a_1$. $a_2' = B - a_1'$, therefore $a_2' < a_2$.

2. If, ceterus paribus, the intervention provides information that the average returns to a_2 are higher than her prior belief, $\mu_2' > \mu_2$, then $a_2' > a_2$. $a_1' = B - a_2'$, therefore $a_1' < a_1$.

3. If the manager is risk-averse ($r > 0$), and ceterus paribus, the intervention provides information that the returns to a_1 are more precise than her prior belief, $\sigma_1' < \sigma_1$, then $a_1' > a_1$. $a_2' = B - a_1'$, therefore $a_2' < a_2$.

4. If the manager is risk-averse ($r > 0$), and ceterus paribus, the intervention provides information that the returns to a_2 are more precise than her prior belief, $\sigma_2' < \sigma_2$, then $a_2' > a_2$. $a_1' = B - a_2'$, therefore $a_1' < a_1$.

Therefore, the intervention, by providing accurate information on Q, allows managers to learn θ_i, $i = \{1, 2\}$, and reallocate resources accordingly, to activities with

[70] Let (26) be written $F(a_1^*(\mu_1, \mu_2, r, \eta, \sigma_1^2, \sigma_2^2), \mu_1, \mu_2, r, \eta, \sigma_1^2, \sigma_2^2) = 0$. Assuming that F is continuous in its arguments and that $\frac{\partial F}{\partial a_1^*} \neq 0$, which follows from the assumptions on g_i, the implicit function theorem can be used to understand the expected responses to new information about the distribution of θ_i

larger average returns or less variable returns to investments. The effect of a simultaneous and equal increase in both μ_1 and μ_2 is ambiguous and depends on the initial value of a_1 and the curvature of g_i. If g_i is linear, then the relative allocation $\frac{a_1}{a_2}$ stays constant. The same applies for simultaneous and equal decreases in σ_1 and σ_2.

Since the optimal allocation of a_1 (a_2) is increasing in μ_1 (μ_2), this means that larger increases in μ_1 (μ_2) will lead to greater increases in a_1 (a_2).[71] Similarly, larger decreases in σ_1 (σ_2) will lead to larger increases in a_1 (a_2). Therefore, the intervention should have a greater impact on outcomes in those areas where managers have less precise information on the returns to activities prior to the intervention – less accurate pre-intervention priors over θ_i. I investigate this margin in the empirical analysis in section 4.4.3.

I have data that provides information on (a_1, a_2) – the final provision of classrooms, and teachers – and data on Q – the enrolment rate, the pupil-class ratio, and the pupil-teacher ratio – and can therefore observe whether there is evidence of reallocation across activities. I do not have data on individual expenditure items to investigate actual monetary investments in the activities, nor the actual returns to these investments. I therefore investigate whether reallocation occurs as a result of the EMIS intervention and 'learning', but I cannot provide direct evidence on exact changes in μ_i or σ_i.

4.4 Empirical analysis

4.4.1 Within-groups panel-data approach

This section describes the identification strategy and empirical approach: a within-groups panel-data approach [Chamberlain, 1984; Wooldridge, 2010]. Since the dataset is of panel form – observing the same districts over time – and the treatment indicator for EMIS training, S_{it}, is changing over time, the following model, under the specified restrictions, identifies the average treatment effect on the treated set of districts.

$$y_{it} = \alpha + \beta S_{it} + \gamma_\tau \mathbf{1}[\tau = t] + \epsilon_{it} \tag{28}$$

$$\epsilon_{it} = f_i + u_{it}$$

$$E(u_{it}|S_{i1}, S_{i2}, ..., S_{iT}, f_i) = 0, \forall\, t = 1, 2, .., T$$

Where y_{it} represents the outcome of interest of district (woreda) i at time t; α the average intercept; γ_τ the aggregate time effect when t is equal to τ; β is the effect of the

[71] The slope of the function $a_1(\mu_1)$ is represented by $\frac{\partial a_1}{\partial \mu_1}$, presented above. The convexity of this function depends on the third derivatives of g_i, which I make no assumptions over.

treatment indicator S_{it}; and ϵ_{it} is the error term, which is comprised of the unobserved time-invariant heterogeneity term, f_i, and the time-varying idiosyncratic error term, u_{it}.

The restriction required for identifying the structural parameter of interest, β, is that: once individual district-level fixed effects, f_i, are controlled for, the treatment indicators, S_{it}, are strictly exogenous with respect to the idiosyncratic error term [Wooldridge, 2010]. This means that all non-random selection into treatment is assumed to occur through f_i and selection into treatment cannot be determined by an idiosyncratic shock. In other words, the treatment and control groups cannot exhibit deterministically different time trends after controlling for the aggregate time trend and the fixed effects.[72] Since the selection into treatment is determined by fixed factors – the distance of the district to the regional capital – the restriction is satisfied.

Table 40 shows the difference-in-differences in the pre-intervention period (between 2007/08 and 2008/09) between those districts treated by the end of the sample timeframe and the 'never-treated' districts to check that the districts selected to receive the intervention show no differential time trends in the pre-intervention period from those that were not selected for treatment (that no idiosyncratic shock in the selected districts determined selection):[73]

Figures 30, 31, and 32 display this graphically with the outcomes of enrolment, pupil-class ratio, and pupil-teacher ratio.[74]

Table 49 shows the same comparison, but for the districts treated in the two big waves of treatment in 2009/10 and 2011/12 versus all other districts, to alleviate concerns about idiosyncratic shocks to these districts determining selection into treatment and finds no evidence for such concerns. Table 50 also provides the regression results of the demographic indicators on the treatment indicator, for whether the district is ever treated in the sample timeframe, to provide an illustration of the fixed differences between the districts (those treated within the sample timeframe tend to be smaller in terms of population, receive smaller budgets, and be poorer).

Since the enrolment rate covers the period 2007/08 to 2012/13 (see table 47), I run the regressions only over this time-period, to observe the changes in other variables using the same timeframe.

The standard errors reported are robust clustered standard errors, clustered at the

[72]This is based on the idea that the control group must form an appropriate counterfactual for the treatment group: that the treatment group would have followed the same trend had it not received treatment. This would be a problem if an individual district's idiosyncratic shock is linked to the decision to enter into treatment or not, à la 'Ashenfelter's dip' [Ashenfelter, 1978].

[73]'never treated' refers to those districts that are not treated at any point during the timeframe of the sample data

[74]The enrolment rate is not available for 2007/08 due to lack of age-specific population data for this year, so I use the nominal number enrolled.

district level, which accounts for arbitrary correlation within groups – serial correlation and heteroskedasticity – but assumes independent errors across districts [Bertrand et al., 2004; Wooldridge, 2010; Cameron and Miller, 2015]. This allows for education outcomes in any particular district to exhibit serial correlation and for the variance of education outcomes within districts to be a function of district characteristics, but does not allow for the unobservable determinants of education outcomes in district i to be related to those in district j; which, for example, rules out cross-district spillovers in shocks that are not accounted for by the aggregate time effects.

4.4.2 Findings

The results, in table 41, show that the training programme leads to a 2.70 percentage-point increase in the enrolment rate (significant at the 0.1% level), meaning that an additional 2.70%, relative to the control group, of all 7-14 year-old children attend primary school as a result of the programme.

Since the national enrolment rate is currently at around 85% of children of 7-14 years attending primary school, the increase represents 18% of the remaining gap towards universal primary education.[75] From the baseline summary statistics in the outcome variables, shown in table 51, we can quantify these effects further: the effect on the enrolment rate represents a 3.2% increase relative to the baseline mean. Using the nominal number of students enrolled in primary schools, as presented in section 4.4.5, implies an additional 167,262 students in primary education due to the programme.[76]

The findings on the other outcome variables, used as measures of the quality of education provision suggest a significant deterioration in the quality of schooling conditions, as the number of pupils increases. Section 4.4.5 provides evidence that the intervention results in significantly fewer teachers and classrooms in the treated districts, relative to the non-treated, as public officials reallocate resources away from these activities and towards enrolment-enhancing activities.

Investigating the dynamic effects of the intervention, I find evidence of short-term declines in the quality of education provision in terms of the quality-of-provision indicators, but evidence of recovery two years after treatment. These results are presented in section 4.4.4.

The cost of implementing the programme is detailed in section 4.7, estimating the costs at between 8,000 USD-16,000 USD per district. A cost-benefit analysis is presented in section 4.4.6.

[75] The drive towards universal primary education is a key government target as well as one of the Millennium Development Goals [United Nations Educational, Scientific and Cultural Organization, 2012; Ministry of Education, Ethiopia, 2015; United Nations, 2015].

[76] 457 multiplied by the 366 treated districts – see table 43

As detailed in section 4.2.6, the mechanism from the new information to service delivery outcomes functions through: hiring and firing decisions, investment and resource-allocation decisions, implementation effort, and organisational management decisions. Due to lack of data on these individual items, the current study is not able to separately identify the contributions of each of these changes to the overall effect.

Section 4.6 provides robustness tests of these results and shows that the results are not driven by cross-sector district-specific factors – the EMIS programme has no significant effects in other areas of service delivery (health and water and sanitation), only in education.

Further exploration into the dynamics of these indicators, as well as an investigation into the heterogeneous treatment effects based on the implications of the conceptual framework follow.

4.4.3 Heterogeneity by the extent of pre-intervention information constraints

This section investigates an implication of the conceptual framework, outlined in section 4.3.2, which states that districts with more severe pre-intervention information constraints should demonstrate greater responses to the intervention. This is explored by estimating the fixed-effects model on different sub-groups of the sample. The sub-groups are defined by the extent to which information is constrained in the **pre-intervention** period, which is proxied by the absolute difference in enrolment numbers in the two most-recent pre-treatment periods (2007/08 and 2009/10). It is assumed that the more that enrolment deviates from period to period, the more severe the information constraints regarding local productivity and the returns to different activities (it is much harder for local bureaucrats to plan effectively and separately identify differential returns to activities from local shocks). Therefore, I expect those with greater fluctuations in enrolment during the pre-intervention period to display larger treatment effects.

The results are presented in figure 33 and show that the impact of the intervention tends to increase with the extent of pre-intervention information constraints faced by the district. The largest coefficient in magnitude is for the most volatile districts and the coefficients tend towards smaller gains in enrolment and schooling conditions as the pre-intervention constraints in information are less severe. Due to the reduction in sample size by breaking the observations into the bins, the differences across percentile groups are suggestive and not statistically significant. These empirical results provide further evidence that information is a key constraint in public service delivery.

4.4.4 Dynamic evidence of multitasking / reallocation

This section explores the existence of a lagged response to understand the effects of the programme over time within each outcome. This question is addressed by using indicators which determine how long ago the district was treated.

The identification conditions are the same as above, but for each of the treatment indicators in table, denoted '$T0_{it}$' to '$T2_{it}$', representing the number of years after that the district has been treated:

$$y_{it} = \alpha + \gamma_\tau \mathbf{1}[\tau = t] + \beta_0 T0_{it} + \beta_1 T1_{it} + \beta_2 T2_{it} + \epsilon_{it}$$

$$\epsilon_{it} = f_i + u_{it}$$

$$E(u_{it}|C_{i1}...C_{iT}, T0_{i1}...T0_{iT}, T1_{i1}...T1_{iT}, T2_{i1}...T2_{iT}, f_i) = 0 \ \ , t = 1, 2, ..., T$$

The omitted base group is 'Never enrolled', C_{it}.

The results, in table 42, show an immediate response in the enrolment rate as a result of the programme. The evidence on the dynamic effects on schooling conditions suggests that the pupil-teacher ratio worsens in the same year as treatment as the number of enrolees increases, though this improves by the first year after the intervention. The pupil-class ratio worsens in the first year after the intervention and improves 2 years after the intervention. Generally, the evidence on the dynamic effects of the programme suggests that districts require two years after the intervention to fully respond in terms of teacher hiring and the provision of classrooms.

4.4.5 Changes in provision: further evidence of reallocation

This section attempts to understand further the response of the treated districts in order to achieve the results in table 41. Table 52 provides the pre-intervention-period difference-in-differences results for this set of outcome variables to show that the common-trends assumption holds for this analysis.

The result in table 43 shows that civil servants in the EMIS districts did indeed generate gains in enrolment, at the expense of significantly fewer teachers and classrooms. Even as the number of pupils increases, creating further pressure on the education system, public officials invest significantly less in the treated districts, relative to the non-treated districts, in terms of the provision of teachers and classes, suggesting a reallocation of resources occurring. Due to the salience of increasing access to education,

public officials are likely to prioritise enrolment-enhancing activities at the expense of other activities to maintain the quality of schooling conditions.[77]

4.4.6 Cost-benefit analysis

Table 43 provides a nominal number for the additional students enrolled per treated district as a result of the programme: 457. Multiplying this by the 366 treated districts gives an additional 167,262 students in primary education as a result of the intervention.

The costs of implementing the programme are detailed in section 4.7 and tables 53 and 54, estimating the costs at between 8,000 USD-16,000 USD per district. I use the upper bound in this section.

The programme, therefore, costs 35 USD per additional student enrolled in the programme. However, this does not take into account costs associated with hiring new teachers and building new classrooms to maintain the quality of education provision.

Using data on the public-sector wage scales, the salary of a newly recruited teacher would be no greater than 1013 ETBm in 2013/14 (the highest salary among the civil service wage scale for first-grade workers). Using this figure, suggests a figure of approximately 60 USD per new teacher recruited.[78] To maintain the pupil-teacher ratio at baseline, as reported in table 51, would need an additional 9 teachers hired for the additional 457 students, per treated district. This translates into 540 USD per treated district.

In addition, according to the fourth Education Sector Development Plan costing from the Ministry of Education, a low-cost primary school classroom is expected to cost 17,700 ETB, which is around 1,000 USD. This means that constructing the 8 classrooms per treated district to maintain the baseline pupil-class ratio would cost 8,000 USD.

Therefore, the total cost of the EMIS programme implementation and the cost of providing additional classrooms and teachers to maintain quality is estimated at 8,981,640 USD.[79]

In this case, the costs per additional pupil enrolled, to maintain the level of inputs per student, is 54 USD.[80] Note that the average expenditure per primary enrolee in 2011/12 (the latest year of available budget data) is 30 USD, suggesting that these

[77]The drive towards universal primary education is a key government target as well as one of the Millennium Development Goals [United Nations Educational, Scientific and Cultural Organization, 2012; Ministry of Education, Ethiopia, 2015; United Nations, 2015].

[78]Exchange rate used is 1 USD to 17.73 ETB, the exchange rate as of 1st July 2012, from xe.com, the last date of the financial year of the sample

[79]This is 16,000 USD per district for the upper-bound cost of the EMIS programme, plus the 540 per treated district to maintain the pupil-teacher ratio, plus the 8,000 per treated district to maintain the pupil-class ratio. The number of treated districts is 366 in the sample.

[80]8,981,640 divided by 167,262.

costs, while small in terms of the potential returns to education, are large in terms of current expenditure levels.

4.5 Mechanisms

The previous analysis begs the question: 'what is the mechanism driving the response from information to provision?' This section explores two potential channels – the use of performance incentives in the office; and the change in the budget allocation and bargaining between regions and districts – neither of which exhibit evidence of driving the response. This section also provides descriptives statistics from the Ethiopia Civil Servants Survey 2016 to explore potential channels further; the descriptive evidence points to informal incentives and career concerns of civil servants as mechanisms for the effect from service delivery information provision to public-service delivery outcomes.

4.5.1 Incentive-based management practices

The dataset used for this paper also contains the responses of the 121 districts identified in both round 3 and round 5 of the Woreda and City Benchmarking Survey (WCBS) and the service delivery dataset [World Bank, 2010*b*]. The survey asks district leaders whether they make use of performance indicators for employees. The use of such an instrument in the workplace is considered as a proxy for the use of incentive-based management practices. It could be that much of the treatment effect is being driven by districts implementing incentive-based management practices and that the effect of information is driven largely through incentives linked to performance. In order to understand the mechanisms from better information to increased effort, the main specification is performed on the WCBS sample, with an additional interaction term between the use of formal incentive-based management practices and the EMIS programme.[81] The results are presented in table 44 and show that formal incentive-based management practices do not seem to be driving the effects of the intervention, since the interaction term is non-positive and statistically insignificant.

4.5.2 Budget reallocation and bargaining between regional governments and district governments

A potential mechanism from information to improved public-service provision could be that the additional information about the level of public services in the district changes the bargaining mechanism between the region and the district for greater budget allocations. The implication would be that the districts that receive the intervention are

[81]The WCBS data used is from the latest round of the Woreda and City Benchmarking Survey 2012/13

better able to bargain for an increase in the budget, due to the availability of accurate information on the status of service delivery outcomes. This mechanism is explored here by running the same analysis as in section 4.4.2, but with the outcome variable as the approved budget of the district. If the programme led to any significant increases in the budgeting process, this would be observed here. I also run the same analysis as section 4.4.2, using a different treatment indicator – whether the district was treated in one of the two big treatment waves – to alleviate the concern that the intervention was accompanied with financial resources. As table 45, shows however, there is no evidence of this, with even negative relative changes in the budget observed for the recipient group. This provides evidence that an increased budget is not driving the results of better education outcomes through the intervention.

4.5.3 Exploring the mechanisms retrospectively through the Ethiopia Civil Servants Survey 2016

Finally, this section uses the evidence gathered in the Ethiopia Civil Servants Survey 2016 (ECSS) to further explore the channels through which the EMIS programme could have had an effect on service provision. The ECSS was carried out across Ethiopia in a representative sample of districts across the five sectors of agriculture, education, health, revenue, and trade offices. Nationally, 2195 civil servants were surveyed, of which 352 were heads of organisation, 465 were directors, and 1366 were employees.[82] Of the 2195, 1436 were district-level civil servants, of which 312 were heads of organisation, 292 were directors, and 828 employees.[83] A total of 66 districts, sampled representatively by region, were surveyed, with the 5 sector offices surveyed within each district. The survey directly asked managers and civil servants in the service about the management-information-system (MIS) programme: how the programme changed the information available to them and affected practices in the workplace; as well as the extent to which information is still a bottleneck to productivity.

The descriptive evidence from the survey relies on recollection, which is subject to bias, and therefore the evidence presented here should be interpreted as suggestive. Nevertheless, the findings provide further food for thought in this area and avenues for future research.

4.5.3.1 The effects of a management information system

The survey asked managers about the effects of the MIS on learning and effort. The questions referred to in this section were asked only to those managers who responded

[82]Out of the 2195, 12 civil servants refused to partake in the survey
[83]At the district level, 4 civil servants out of the 1436 refused to partake

affirmatively to a MIS being in place and that were in the organisation when the MIS was introduced, across all three tiers of government, across all five sectors included in the survey.

The tables for this section are provided in the appendix. Table 55 and 56 confirm that around 85% of managers felt that the MIS led explicitly to learning new information on the state of service delivery and the conditions in the jurisdiction.

Table 57 suggests that the MIS often (82% of managers agree to some extent) does feed into performance reviews; this information contrasts with the results in section 4.5.1. Together, this implies that the information provided by the EMIS indicators feeds into (subjective) performance reviews, which incentivises the civil servants through career concerns (i.e. promotion prospects and future career tracks) and non-monetary incentives (such as training opportunities, out-of-country visits, recognition awards, and attractive placements), rather than through the monetary incentives tied to individual formal performance indicators. This descriptive evidence, therefore, suggests that the use of the service delivery information in (subjective) performance reviews interacts with career concerns and non-formal incentives to generate additional effort from civil servants.

4.5.3.2 Current information constraints in the service

Tables 58 and 59 refer to questions asked to all employees, across all three tiers of government, across all five sectors included in the survey, regarding their perceptions of the availability of information regarding service delivery and jurisdiction conditions.

In terms of the constraints on information, table 58 shows that from the 1324 employees (excluding those that replied 'Don't know' and 'Refused to answer'), the average percentage of projects for which employees feel they have sufficient information on the state of service delivery in the jurisdiction is 71%, while this is 60% when asked about information on other conditions in the jurisdiction (table 59). This suggests that the greatest information constraint to civil servants in the Ethiopian service is information on district conditions, such as demographic, social, and economic indicators. Therefore, understanding further the potential consequences of marginal improvements in different types of information to civil servants is clearly another area for future research in this area.

4.6 Robustness

4.6.1 Do the effects show up in other public service delivery outcomes?

A test of robustness is to check that the effects of the treatment are not being driven by something more systematically related to district public service delivery performance, but rather directly by the evaluated programme. To check this, table 46, shows the results from the main specification on the following outcomes: the contraceptive acceptance rate, the rate of births delivered by a skilled attendant, the vaccination rate, and the number of public water supply points in the jurisdiction.[84] If the treatment was having an impact on other public service delivery outcomes, this would suggest an alternative explanation to the main results than that through changes occurring in the education offices. The results show no effect of the treatment on the other outcomes, providing further convincing support for the identification strategy and main findings of this paper.

4.7 Conclusions

This paper provides evidence that public servants face information constraints that affect their performance. Indeed, I provide evidence that an e-governance intervention in the Ethiopian bureaucracy leads to a 2.7 percentage-point increase in the enrolment of 7-14 year-old children in primary school on average, which is 18% of the remaining gap towards universal primary education. The results on the number of enrolees puts the figure at an additional 167,262 students in primary education as a result of the intervention.

There is evidence of a decline in the quality of education provision as a result of the expansion in enrolment. This is in line with the new information from the management-information system leading to a reallocation of resources across activities. The decline in quality begins to recover after the second year after treatment, as districts hire more teachers and create more teaching space.

The empirical evidence also supports an additional implication of the conceptual framework: those districts facing greater pre-intervention constraints to information display larger responses as a result of the intervention, suggesting that information is indeed a constraint to public service delivery.

The analysis provides evidence that the use of formal incentive-based management practices and budget-bargaining considerations between the regional and district gov-

[84]The number of public water supply points is from the Woreda and City Benchmarking Survey and hence covers only the 114 districts in the matched WCBS sample, for both rounds of the survey (2009/10 and 2012/13), for which there is data on the water points. The vaccination rate refers to the vaccination of Penta-3.

ernments are not driving the responses to the intervention. Descriptive statistics from the Ethiopia Civil Servants Survey 2016 provide suggestive evidence that the channels through which the effects are functioning are through non-monetary incentives and career concerns.

The programme passes any reasonable cost-benefit evaluation, requiring low returns to education to cover the 54-USD-per-additional-enrolee cost, accounting for costs related to hiring new teachers and building new classrooms to maintain schooling conditions. These costs, however, are large relative to current levels of per-pupil expenditure (30 USD).

The results are robust to checks against endogenous selection into the intervention related to unobserved locality-specific factors.

This paper contributes to the research on the productivity of public-sector organisations, specifically investigating the role of information provision to managers and civil servants in public-sector organisations, as well as providing a further understanding of the benefits of e-governance initiatives and their cost-effectiveness.

Appendix: Costs

This section provides an outline of the costs of the programme.

Annual work plans from the General Education Quality Improvement Programme of the Ministry of Education allow upper and lower bounds of the cost of establishing an effectively-running EMIS at the district level to be calculated [World Bank, 2015*c*].[85] The upper bound is calculated using the national-level cost information of establishing a fully effective EMIS programme, which is decentralised to regions and districts; this cost, therefore, includes the costs of federal-level, regional-level, and district-level development of EMIS. This cost is considered due to the fact that an effective running of district-level EMIS relies on functioning EMIS at higher levels in order to provide support and direction to the district – for example, to provide the questionnaires, the software, the training, to help consolidate regional- and national-level data, and occasionally personnel. The lower bound cost takes the isolated cost of extending the functioning of EMIS down to district level, conditional on a functional EMIS at federal- and regional-level. This lower bound is interesting as it provides an isolated view of the programme's cost, conditional on all other factors being in place for the programme to be implemented. Though there will be additional recurring costs, due to turnover of personnel, institutional memory constraints, and human and physical capital replacement, these are ignored in order to consider the cost and benefits of the programme in

[85] http://www.worldbank.org/projects/P129828/ethiopia-general-education-quality-improvement-project-ii?lang=en [Accessed 31 July 2018]

absence of these additional frictions.

The upper-bound cost is presented in table 53; the lower-bound cost in table 54. The per-district costs are based on the total number of districts treated in the sample timeframe, 475.

Table 39: Intervention rollout

	No EMIS		EMIS	
	Frequency	Percent	Frequency	Percent
2006/07	503	100.00	0	0.00
2007/08	503	100.00	0	0.00
2008/09	503	100.00	0	0.00
2009/10	380	75.55	123	24.45
2010/11	380	75.55	123	24.45
2011/12	137	27.24	366	72.76
2012/13	137	27.24	366	72.76
2013/14	137	27.24	366	72.76

Notes: Frequencies and row percentages of districts with and without the intervention. The sample is the full set of districts used for the analysis, for the entire timeframe of the dataset. The Ministry of Education database has a total of 707 districts outside of Addis Ababa[a]. The database used in this analysis, based on data coverage for all indicators for all years, contains 503 districts. Table 48 compares the observable characteristics in the pre-treatment year of 2008/09 of those in the database and those not in the database. Figures are rounded to 2 decimal places.

[a]The 2007 Ethiopian census contains 726 districts excluding special administrative areas and Addis Ababa.

Table 40: Checking for common trends in the pre-treatment period

	Enrolment rate	Pupil-teacher ratio	Pupil-class ratio
Treated in sample timeframe x 1[2008/9]	1.77	-0.07	0.46
	(1.46)	(0.71)	(0.58)
Year fixed effects	Yes	Yes	Yes
Adjusted R-Squared	0.98	0.91	0.92
Observations	1006	1006	1006
Clusters	503	503	503

Notes: Significance levels: * p<0.1, ** p<0.05, *** p<0.01. Standard errors in parentheses. Standard errors are clustered at the district level. The unit of observation is the district. The sample is restricted to the pre-treatment period of 2007/8 to 2008/9. The table displays coefficients and standard errors from an OLS regression of the dependent variable on a set of year fixed effects, district fixed effects, and the interaction between the year effect and the treatment indicator. Treated in sample timeframe is an indicator equal to one if the district receive the EMIS programme at any point within the sample timeframe; 1[2008/9] is an indicator for whether the year is 2008/9. The columns refer to the different outcome variables used in each specification. See table 47 for further details on the variables. The results represent the difference-in-differences comparison of the treated versus the never-treated groups in the pre-treatment years of 2007/08 and 2008/09. Figures are rounded to 2 decimal places.

Table 41: Average effects of the EMIS programme on education provision

	Enrolment rate	Pupil-teacher ratio	Pupil-class ratio
EMIS programme [Treatment=1]	2.70***	3.56	1.10***
	(0.71)	(3.34)	(0.39)
Year fixed effects	Yes	Yes	Yes
Adjusted R-Squared	0.97	0.20	0.84
Observations	3018	3018	3018
Clusters	503	503	503

Notes: Significance levels: * p<0.1, ** p<0.05, *** p<0.01. Standard errors in parentheses. Standard errors are clustered at the district level. The unit of observation is the district. The table displays coefficients and standard errors from an OLS regression of the dependent variable on a set of year fixed effects, district fixed effects, and the time-varying treatment indicator for whether the district has an EMIS in place. The columns refer to the different outcome variables used in each specification. See table 47 for further details on the variables. Figures are rounded to 2 decimal places.

Table 42: Average effects of the EMIS programme on education provision, by duration in the programme

	Enrolment rate	Pupil-teacher ratio	Pupil-class ratio
Treated this year	3.01***	4.99	-0.11
	(0.72)	(5.78)	(0.46)
Treated 1 year ago	-0.60	-3.77	2.19***
	(0.58)	(5.50)	(0.52)
Treated 2 years ago	0.40	-3.68***	-2.50***
	(1.18)	(0.87)	(0.77)
Year fixed effects	Yes	Yes	Yes
Adjusted R-Squared	0.97	0.20	0.85
Observations	3018	3018	3018
Clusters	503	503	503

Notes: Significance levels: * p<0.1, ** p<0.05, *** p<0.01. Standard errors in parentheses. Standard errors are clustered at the district level. The unit of observation is the district. The table displays coefficients and standard errors from an OLS regression of the dependent variable on a set of year fixed effects, district fixed effects, and a set of dummy variables indicating when the district received the EMIS programme. The columns refer to the different outcome variables used in each specification. See table 47 for further details on the variables. Figures are rounded to 2 decimal places.

Table 43: Average effects of the EMIS programme on education provision

	Enrolment	Teachers	Classes
EMIS programme [Treatment=1]	456.93**	-25.29***	-14.77***
	(218.22)	(8.02)	(3.24)
Year fixed effects	Yes	Yes	Yes
Adjusted R-Squared	0.97	0.78	0.97
Observations	3018	3018	3018
Clusters	503	503	503

Notes: Significance levels: * p<0.1, ** p<0.05, *** p<0.01. Standard errors in parentheses. Standard errors are clustered at the district level. The unit of observation is the district. The table displays coefficients and standard errors from an OLS regression of the dependent variable on a set of year fixed effects, district fixed effects, and the time-varying treatment indicator for whether the district has an EMIS in place. The columns refer to the different outcome variables used in each specification. See table 47 for further details on the variables. Figures are rounded to 2 decimal places.

Table 44: Are the treatment effects larger when incentive-based management practices are in place?

	Enrolment rate	Pupil-teacher ratio	Pupil-class ratio
EMIS programme [Treatment=1]	1.81	3.06	1.75
	(1.37)	(5.25)	(1.11)
EMIS programme x [Performance indicators=1]	1.26	7.28	-1.75
	(2.49)	(9.69)	(1.50)
Year fixed effects	Yes	Yes	Yes
Year x Performance indicator fixed effects	Yes	Yes	Yes
Adjusted R-Squared	0.98	0.30	0.86
Observations	726	726	726
Clusters	121	121	121

Notes: Significance levels: * p<0.1, ** p<0.05, *** p<0.01. Standard errors in parentheses. Standard errors are clustered at the district level. The unit of observation is the district. The sample used in the analysis is the set of districts for which there is institutional information from the Woreda and City Benchmarking Survey. The table displays coefficients and standard errors from an OLS regression of the dependent variable on a set of year fixed effects, a set of year fixed effects interacted with a dummy variable indicating whether the district uses formal performance indicators to incentivise employees, district fixed effects, the time-varying treatment indicator for whether the district has an EMIS in place, and an interaction term between the use of formal performance indicators in the district and the time-varying treatment indicator. The columns refer to the different outcome variables used in each specification. See table 47 for further details on the variables. Figures are rounded to 2 decimal places.

Table 45: Is the response driven by increases in the budget?

	Total Approved Budget (ETBm)	
	Intervention	Treated in 2009/10 or 11/12
EMIS programme [Treatment=1]	-1.13*	
	(0.62)	
Treated in 2009/10 or 2011/12		-2.79***
		(0.77)
Year fixed effects	Yes	Yes
Adjusted R-Squared	0.84	0.85
Observations	1988	1988
Clusters	497	497

Notes: Significance levels: * p<0.1, ** p<0.05, *** p<0.01. Standard errors in parentheses. Standard errors are clustered at the district level. The unit of observation is the district. The table displays coefficients and standard errors from an OLS regression of the dependent variable on a set of year fixed effects, district fixed effects. [Treatment=1] is the time-varying treatment indicator for whether the district has an EMIS in place. Treated in 2009/10 or 2011/12 is an indicator equal to one if the district was treated in the two main waves. The dependent variable is the approved budget of the district in millions of Ethiopian Birr, sourced from the Open Budget Portal initiative of the World Bank. The coverage of this variable is between 2008/09 and 2011/12. See table 47 for further details on the variables. Figures are rounded to 2 decimal places.

Table 46: Does the treatment have an effect on health and water service outcomes?

	Contraceptive acceptance rate	Rate of births delivered by skilled attendant	Vaccination rate	No. of public water supply points
EMIS programme [Treatment=1]	-1.95	1.38	-2.30	-47.62
	(2.09)	(1.69)	(1.63)	(39.50)
Year fixed effects	Yes	Yes	Yes	Yes
Adjusted R-Squared	0.45	0.28	0.43	0.35
Observations	1624	1551	1570	190
Clusters	503	503	503	114

Notes: Significance levels: * p<0.1, ** p<0.05, *** p<0.01. Standard errors in parentheses. Standard errors are clustered at the district level. The unit of observation is the district. The table displays coefficients and standard errors from an OLS regression of the dependent variable on a set of year fixed effects, district fixed effects, and the time-varying treatment indicator for whether the district has an EMIS in place. The columns refer to the different outcome variables used in each specification. The Contraceptive Acceptance Rate, Rate of births delivered by a skilled attendant, and Vaccination rate are from the Ministry of Health Core Plans and cover the periods from 2008/09 to 2011/12; the Number of public water supply points is from the Woreda and City Benchmarking Survey for the years 2009/10 and 2012/13, the two rounds of the survey. See table 47 for further details on the variables. Figures are rounded to 2 decimal places.

Table 47: Variables used in the analysis and the coverage

Indicator	Description	Source	Coverage
Enrolment	The enrolment in the district for grades 1-8 (primary school), male and female	Ministry of Education, Ethiopia	2006/07-2013/14
Classes	The sections in the district for grades 1-8 (primary school), male and female	Ministry of Education, Ethiopia	2006/07-2013/14
Teachers	The number of teachers in the district for grades 1-8 (primary school), male and female	Ministry of Education, Ethiopia	2006/07-2012/13
Pupil-class ratio	The pupil-section ratio in the district for grades 1-8 (primary school), male and female	Ministry of Education, Ethiopia	2006/07-2013/14
Pupil-teacher ratio	The pupil-teacher ratio in the district for grades 1-8 (primary school), male and female	Ministry of Education, Ethiopia	2006/07-2012/13
Enrolment rate	The enrolment rate (the total number of correct-age enrollees over the population of correct-age children) in the district for grades 1-8 (primary school), male and female	Ministry of Education, Ethiopia	2007/08-2012/13
EMIS programme	Indicates whether the district received training on the EMIS database software by the respective year	Ministry of Education, Ethiopia	2009/10-2013/14
Woreda budget transfer	Annual budget transfer to the district	World Bank Open Budget Portal	2008/09-2011/12
Contraceptive acceptance rate	Percentage of population in the district that is using contraceptive	World Bank Ethiopia Poverty and Social Impact Analysis Database	2008/09-2011/12
Rate of births delivered by a skilled attendant	Percentage of births that are delivered by a skilled health personnel	World Bank Ethiopia Poverty and Social Impact Analysis Database	2008/09/2011-12
Vaccination rate	Percentage of population that have received the Penta-3 vaccine	World Bank Ethiopia Poverty and Social Impact Analysis Database	2008/09-2011/12
Number of public water supply points	The number of public water supply points in the jurisdiction	Woreda and City Benchmarking Survey	2009/10 and 2012/13
Percentage of rural inhabitants	Percentage of the district's population that is considered rural	Census 2007	2007/08
Population	District population from the census	Census 2007	2007/08
Poverty rate	District-level poverty rate from the census	Census 2007	2007/08

Notes: Enrolment, classes, pupils per classes, pupils per teacher, and enrolment rate all sourced directly from the Ministry of Education, Ethiopia, statistics database onsite. The indicator for whether the EMIS is in place is sourced directly from the Ministry of Education, Ethiopia, from a survey conducted specifically for the purpose of this study. The district budget transfer data is sourced from the Open Budget Portal initiative of the World Bank[a]. The percentage of rural inhabitants, the population, and the poverty rate are all data from the 2007 Ethiopian census[b]. The poverty rate used is based on the national poverty line[c]; for 2011, the national poverty line and the poverty line using $1.25USD (PPP) were almost identical (30% living below the poverty line and 31% living below the poverty line respectively) [World Bank, 2015*a*]. Health data are taken from World Bank Ethiopia's Poverty and Social Impact Analysis Database [Khan et al., 2014]. Water supply data is sourced from the Woreda and City Benchmarking Survey [World Bank, 2010*b*].

[a] http://wbi.worldbank.org/boost/boost-initiative [Accessed 31 July 2018]

[b] http://www.csa.gov.et/census-report/complete-report/census-2007 [Accessed 31 July 2018]

[c] This combines household consumption expenditure data from the Household Income, Consumption and Expenditure Survey with the census data

Table 48: Representativeness of sampled districts: observable characteristics in 2008/09

	Mean	St Dev	Median	N
Not all years				
Primary enrolment rate	65.97	42.43	71.59	152
Primary pupil-teacher ratio	50.28	16.29	47.85	151
Primary pupil-class ratio	76.16	55.57	57.85	203
Population, millions	0.09	0.06	0.08	204
Percentage of rural inhabitants	82.19	27.54	91.94	204
Poverty rate	28.81	14.40	26.29	204
District budget in ETB millions	14.39	13.23	12.58	199
All years of data				
Primary enrolment rate	85.37	49.68	82.11	503
Primary pupil-teacher ratio	54.92	14.34	53.90	503
Primary pupil-class ratio	59.37	12.15	57.67	503
Population, millions	0.10	0.06	0.09	502
Percentage of rural inhabitants	84.92	25.16	92.95	503
Poverty rate	29.80	12.84	27.74	502
District budget in ETB millions	15.37	7.19	13.89	497
Total				
Primary enrolment rate	80.87	48.76	81.03	655
Primary pupil-teacher ratio	53.85	14.93	52.47	654
Primary pupil-class ratio	64.19	32.37	57.73	706
Population, millions	0.10	0.06	0.09	706
Percentage of rural inhabitants	84.13	25.88	92.52	707
Poverty rate	29.52	13.30	27.61	706
District budget in ETB millions	15.09	9.33	13.67	696

Notes: The table provides summary statistics of the sampled and non-sampled districts for the pre-treatment year of 2008/09. The unit of observation is the district. The sampled districts are those that are observed for all years of the data. The rows refer to the variable and the column to the statistic. See table 47 for details on the sources and coverage of each indicator.

Table 49: Checking for common trends in the pre-treatment period

	Enrolment rate	Pupil-teacher ratio	Pupil-class ratio
Treated in 2009/10 or 2011/12 x 1[2008/9]	1.77	-0.07	0.46
	(1.46)	(0.71)	(0.58)
Year fixed effects	Yes	Yes	Yes
Adjusted R-Squared	0.98	0.91	0.92
Observations	1006	1006	1006
Clusters	503	503	503

Notes: Significance levels: * p<0.1, ** p<0.05, *** p<0.01. Standard errors in parentheses. Standard errors are clustered at the district level. The unit of observation is the district. The sample is restricted to the pre-treatment period of 2007/8 to 2008/9. The table displays coefficients and standard errors from an OLS regression of the dependent variable on a set of year fixed effects, district fixed effects, and the interaction between the year effect and the treatment indicator. Treated in 2009/10 or 2011/12 is an indicator equal to one if the district received the EMIS programme during one of the two big waves of treatment; 1[2008/9] is an indicator for whether the year is 2008/9. The columns refer to the different outcome variables used in each specification. The results represent the difference-in-differences comparison of the treated-in-2009/10 or treated-in-2011/12 (the two large waves of treatment) versus the other districts in the pre-treatment years of 2007/08 and 2008/09. Figures are rounded to 2 decimal places.

Table 50: Pre-treatment differences in districts across demographics

	Budget ETB millions	Population millions	Poverty rate	Percentage rural
Treated in sample timeframe	-8.13***	-0.03***	3.85***	-0.00
	(0.84)	(0.01)	(1.26)	(0.02)
Adjusted R-Squared	0.25	0.048	0.016	-0.0020
Observations	497	502	502	503
Clusters	497	502	502	503

Notes: Significance levels: * p<0.1, ** p<0.05, *** p<0.01. Standard errors in parentheses. Standard errors are clustered at the district level. The unit of observation is the district. The sample is restricted to the year 2008/09, the last year prior to treatment. The table displays coefficients and standard errors from an OLS regression of the dependent variable on an indicator for whether the district is ever treated in the sample. Treated in sample timeframe is an indicator equal to one if the district received the EMIS programme at any point during the sample timeframe. The columns refer to the different outcome variables used in each specification. The results represent the pre-treatment comparison of the treated versus the never-treated groups in the pre-treatment years of 2007/08 and 2008/09. Figures are rounded to 2 decimal places.

Table 51: Summary statistics of outcome variables in 2008/09: mean and standard deviation

	Mean	St Dev	Median	N
Primary enrolment rate	85.37	49.68	82.11	503
Primary pupil-teacher ratio	54.92	14.34	53.90	503
Primary pupil-class ratio	59.37	12.15	57.67	503

Notes: The table shows means and standard deviations for the main outcome variables of interest, shown in the rows of the table, in the year prior to treatment (2008/09). The rows refer to the variable and the column to the statistic. The unit of observation is the district. Treatment began being rolled out in 2009/10. See table 47 for further details of the variables. Figures are rounded to 2 decimal places.

Table 52: Checking for common trends in the pre-treatment period

	Enrolment	Teachers	Classes
Treated in sample timeframe x 1[2008/9]	449.65	-8.93	-1.10*
	(355.83)	(6.03)	(0.64)
Year fixed effects	Yes	Yes	Yes
Adjusted R-Squared	0.99	0.98	0.97
Observations	1006	1006	1006
Clusters	503	503	503

Notes: Significance levels: * p<0.1, ** p<0.05, *** p<0.01. Standard errors in parentheses. Standard errors are clustered at the district level. The unit of observation is the district. The sample is restricted to the pre-treatment period of 2007/8 to 2008/9. The table displays coefficients and standard errors from an OLS regression of the dependent variable on a set of year fixed effects, district fixed effects, and the interaction between the year effect and the treatment indicator. Treated in sample timeframe is an indicator equal to one if the district received the EMIS programme at any point during the sample timeframe; 1[2008/9] is an indicator for whether the year is 2008/9. The columns refer to the different outcome variables used in each specification. The results represent the difference-in-differences comparison of the treated versus the never-treated groups in the pre-treatment years of 2007/08 and 2008/09. Figures are rounded to 2 decimal places.

Table 53: Upper bound costs based on GEQIP Information of EMIS development at regional and district level

Cost	Amount in ETB	Amount in USD	Amount per district
Support the decentralisation of EMIS at the regional and district level.	731,000 ETB	41,230 USD	112.65 USD
Orient preliminary data collection formats and capacity building to REBs on the software.	211,748 ETB	11,943 USD	32.63 USD
To train MoE and REB EMIS experts on technology use, software management, analysis, questionnaire development, maintenance, networking.	3,687,892 ETB	208,003 USD	568.31 USD
Training one EMIS facilitator at each district.	10,939,500 ETB	617,005 USD	1685.81 USD
Additional support to emerging regions to strength EMIS.	3,603,619ETB	203,250 USD	555.33 USD
Informing data users on the education indicators and data use.	5,265,323 ETB	296,973 USD	811.40 USD
Develop data management at the district level.	1,200,000 ETB	67,682 USD	184.92 USD
Printing and development of manuals and guidelines for EMIS implementation.	1,697,500 ETB	95,742 USD	261.59 USD
Developing and customising the EMIS software.	2,780,000 ETB	156,796 USD	428.41 USD
Internet connectivity.	3,992,000 ETB	225,155 USD	615.18 USD
Initial procurement of EMIS hardware for all districts.	67,331,708 ETB	3,797,615 USD	10,376.00 USD
Total	101,440,290 ETB	5,721,393 USD	15,632 USD

Notes: The table shows the cost estimates of the implementation of the entire EMIS programme. The upper bound is calculated using the national-level cost information of establishing a fully effective EMIS programme, which is decentralised to regions and districts; this cost, therefore, includes the costs of federal-level, regional-level, and district-level development of EMIS. The costs are gathered from the annual work-plans of the General Education Quality Improvement Programme (GEQIP): http://www.worldbank.org/projects/P129828/ethiopia-general-education-quality-improvement-project-ii?lang=en [Accessed 31 July 2018]. Exchange rate used is 1 USD to 17.73 ETB, the exchange rate as of 1st July 2012, from xe.com, the last date of the financial year of the sample. The per-district costs take the 366 districts that are treated in the sample.

Table 54: Lower bound costs based on GEQIP Information of district EMIS training, procurement, and data management

Cost	Amount in ETB	Amount in USD	Amount per district
Training one EMIS facilitator at each district.	10,939,500 ETB	617,005 USD	1,685.81 USD
Develop data management at the district level.	1,200,000 ETB	67,682 USD	184.92 USD
Initial procurement of EMIS hardware for all districts.	67,331,708 ETB	3,797,615 USD	10,376.00 USD
Total	79,471,208 ETB	3,824,961 USD	8,052.55 USD

Notes: The table shows the cost estimates of the implementation of the entire EMIS programme. The lower bound cost takes the isolated cost of extending the functioning of EMIS down to district level, conditional on a functional EMIS at federal- and regional-level. The costs are gathered from the annual work-plans of the General Education Quality Improvement Programme (GEQIP): http://www.worldbank.org/projects/P129828/ethiopia-general-education-quality-improvement-project-ii?lang=en [Accessed 31 July 2018]. Exchange rate used is 1 USD to 17.73 ETB, the exchange rate as of 1st July 2012, from xe.com, the last date of the financial year of the sample. The per-district costs take the 366 districts that are treated in the sample.

Table 55: The introduction of the MIS increased the information that I had access to regarding the state of service delivery in the jurisdiction

	Frequency	Percent
Strongly disagree	9	6.16
Slightly disagree	8	5.48
Neutral	5	3.42
Slightly agree	61	41.78
Strongly agree	63	43.15
Total	146	100.00

Notes: The table shows the frequencies of the responses to the survey item 'To what extent do you agree with the following statements? The introduction of the MIS increased the information that I had access to regarding the state of service delivery in the jurisdiction'. This item was asked to directors only; and only those directors that were in the organisation when the MIS was introduced; and only directors from organisations that have a functioning MIS in place.

Table 56: I learnt new things about the jurisdiction that I never knew before because of the MIS

	Frequency	Percent
Strongly disagree	12	8.22
Slightly disagree	6	4.11
Neutral	5	3.42
Slightly agree	64	43.84
Strongly agree	57	39.04
Do not know	2	1.37
Total	146	100.00

Notes: The table shows the frequencies of the responses to the survey item 'To what extent do you agree with the following statements? I learnt new things about the jurisdiction that I never knew before because of the MIS'. This item was asked to directors only; and only those directors that were in the organisation when the MIS was introduced; and only directors from organisations that have a functioning MIS in place.

Table 57: The information provided / generated by the MIS feeds into the performance review process or another accountability mechanism

	Frequency	Percent
Strongly disagree	12	8.22
Slightly disagree	6	4.11
Neutral	4	2.74
Slightly agree	72	49.32
Strongly agree	48	32.88
Not applicable	1	0.68
Do not know	3	2.05
Total	146	100.00

Notes: The table shows the frequencies of the responses to the survey item 'To what extent do you agree with the following statements? The information provided / generated by the MIS feeds into the performance review process or another accountability mechanism'. This item was asked to directors only; and only those directors that were in the organisation when the MIS was introduced; and only directors from organisations that have a functioning MIS in place.

Table 58: I have necessary information regarding the current state of service delivery in my jurisdiction/my area of work

	Mean	St Dev	Min	Max
Necesssary information on service delivery	70.92	19.12	0.00	100.00
Observations	1324			

Notes: The table shows the summary statistics of the responses to the survey item 'In what proportion of projects/programmes that you work on would you say that the following statements were true: I have necessary information regarding the current state of service delivery in my jurisdiction/my area of work'. This item was asked to employees only. 1366 employees were asked this question; the above excluded those that replied 'Don't know' or 'Refused to answer".

Table 59: I have necessary information regarding other conditions in the jurisdiction/my area of work, such as the level of poverty, the level of education/literacy, the demographics, etc

	Mean	St Dev	Min	Max
Necesssary information on district conditions	59.80	23.44	0.00	100.00
Observations	1286			

Notes: The table shows the summary statistics of the responses to the survey item 'In what proportion of projects/programmes that you work on would you say that the following statements were true: I have necessary information regarding other conditions in the jurisdiction/my area of work, such as the level of poverty, the level of education/literacy, the demographics, etc'. This item was asked to employees only. 1366 employees were asked this question; the above excluded those that replied 'Don't know' or 'Refused to answer".

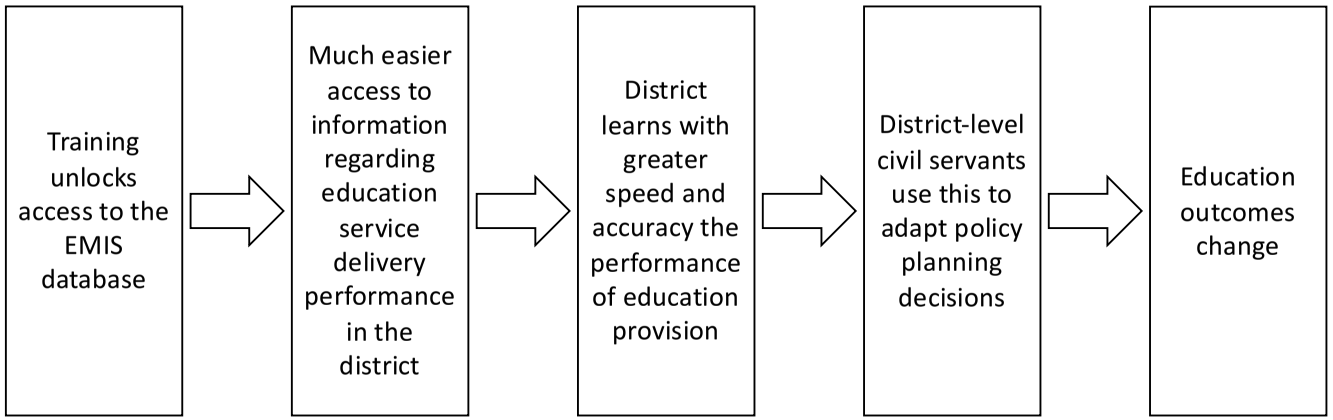
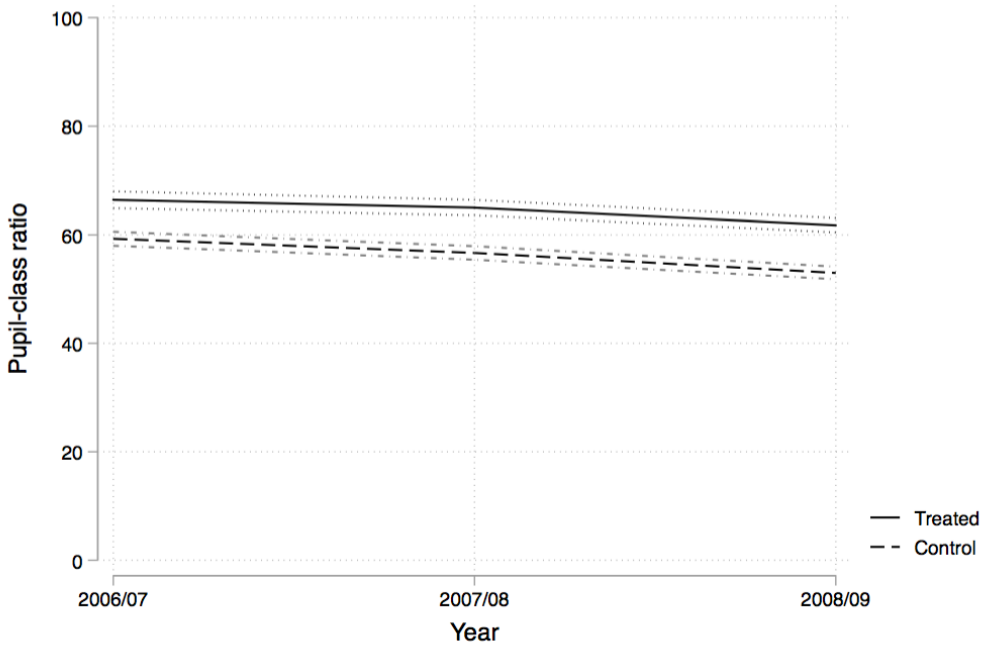
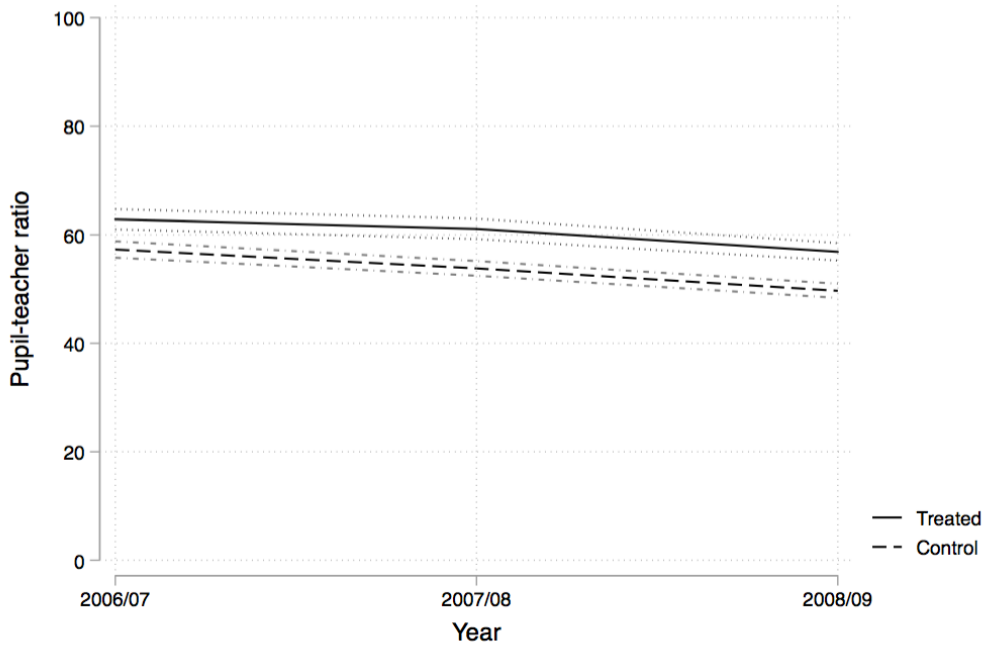


Figure 29: Channels of effects from the training programme



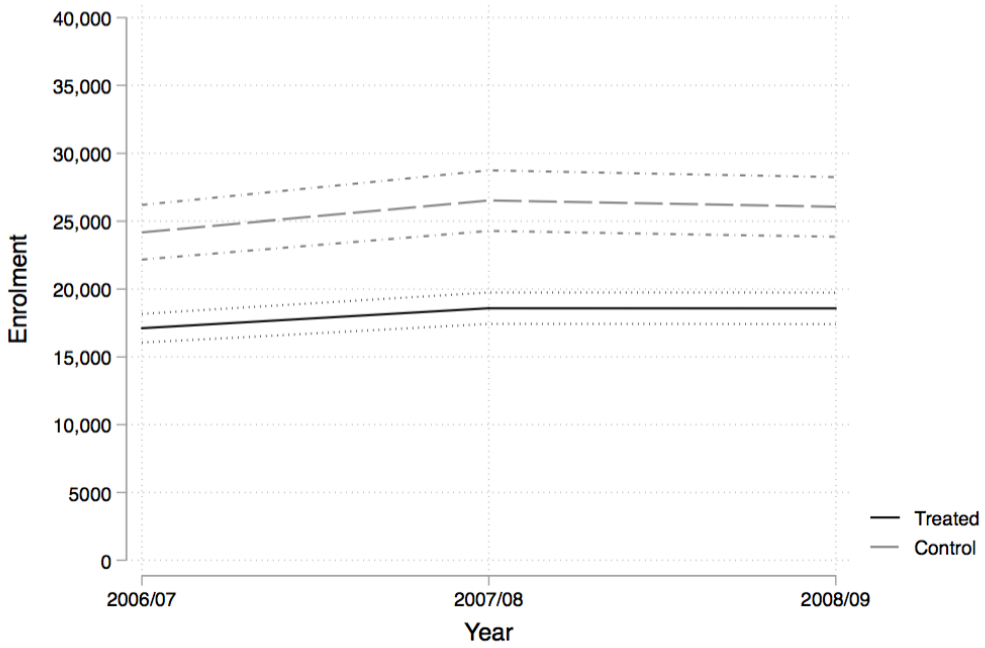
Notes: The figure shows the mean pupil-class ratio for the treated districts (black solid line) and 95% confidence intervals about the mean (grey dotted lines). The figure also shows the mean pupil-class ratio for the control districts (black dashed line) and 95% confidence intervals about the mean (grey dot-dashed lines). 'Treated' districts refers to those districts that are treated at some point during the sample timeframe. 'Control' districts refer to those that are never treated during the sample timeframe. The unit of observation is the district. The timeframe shown in the figure covers the entire pre-treatment period within the sample dataset timeframe.

Figure 30: Pre-treatment trend: pupil-class ratio



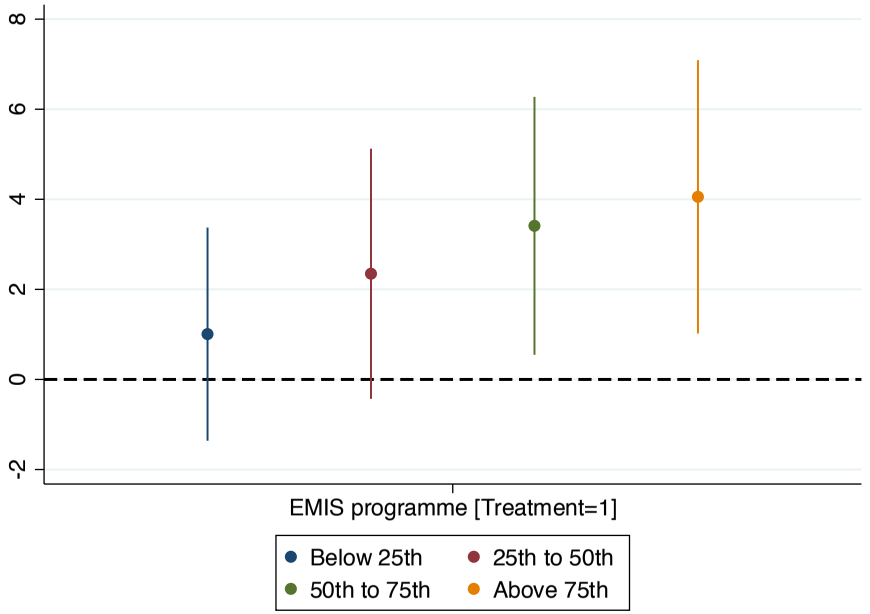
Notes: The figure shows the mean pupil-teacher ratio for the treated districts (black solid line) and 95% confidence intervals about the mean (grey dotted lines). The figure also shows the mean pupil-teacher ratio for the control districts (black dashed line) and 95% confidence intervals about the mean (grey dot-dashed lines). 'Treated' districts refers to those districts that are treated at some point during the sample timeframe. 'Control' districts refer to those that are never treated during the sample timeframe. The unit of observation is the district. The timeframe shown in the figure covers the entire pre-treatment period within the sample dataset timeframe.

Figure 31: Pre-treatment trend: pupil-teacher ratio



Notes: The figure shows the mean enrolment for the treated districts (black solid line) and 95% confidence intervals about the mean (grey dotted lines). The figure also shows the mean enrolment for the control districts (black dashed line) and 95% confidence intervals about the mean (grey dot-dashed lines). 'Treated' districts refers to those districts that are treated at some point during the sample timeframe. 'Control' districts refer to those that are never treated during the sample timeframe. The unit of observation is the district. The timeframe shown in the figure covers the entire pre-treatment period within the sample dataset timeframe.

Figure 32: Pre-treatment trend: enrolment



Notes: The figure displays the coefficient on a time-varying treatment indicator for whether the district has an EMIS in place from an OLS regression of the dependent variable on a set of year fixed effects, district fixed effects, and the treatment indicator. The 95% confidence intervals are also shown, where standard errors are clustered at the district level. The unit of observation is the district. The dependent variable for the regression is the enrolment rate in the district. The regression is conducted on four sub-samples, related to the pre-treatment-period volatility in the enrolment of students for the district. The pre-treatment-period volatility is the absolute difference in enrolment numbers in the two most-recent pre-treatment periods (2007/08 and 2008/09). The columns refer to the different sub-samples of districts: below 25th is the sub-sample of districts in which the absolute difference in pre-treatment enrolment is in the 25th percentile or lower; 25th to 50th the districts with a pre-treatment absolute difference in enrolment between the 25th and 50th percentiles, and so on. Figures are rounded to 2 decimal places. See table 47 for further details on the variables.

Figure 33: Average effects of the EMIS programme on the enrolment rate, by quantile of pre-treatment-period information constraints

5 Efficiency Losses from Subjective Performance Evaluations and Supervisor Bias

5.1 Introduction

Standard agency theory studies optimal incentives in the presence of objective performance measures. The principal, who is the residual claimant of output, evaluates the performance of the agent and transfers the appropriate compensation [Hart and Holmstrom, 1987]. Most compensation arrangements depart from this in two ways: (i) the evaluation is conducted by a a middle-tier supervisor, who is not the residual claimant of output; and (ii) the evaluation is based on subjective judgements of performance rather than objective measures.[86] These two deviations allow for supervisor bias to creep into performance evaluations, potentially leading to efficiency losses within the organisation.[87] Despite the prevalence of subjective performance evaluations within organisations, little empirical research has been conducted in this area, likely due to the data limitations [Frederiksen et al., 2017].[88]

Using new data on individual performance-appraisal data from the Ethiopian bureaucracy on objective and subjective performance measures, this paper provides the first attempt, to my knowledge, to put an explicit value on supervisor bias at the individual worker level and correlate this to employee wages. In addition, this paper provides a measure of the cost of supervisor bias at the organisation level in terms of service delivery performance. The dual-reporting system used in the Ethiopian bureaucracy allows me to construct a measure for supervisor bias. The Balanced Scorecard system of the Ethiopian bureaucracy uses a performance score and an 'attitude' score to develop an overall appraisal score for each employee. The performance score is based on a weighted average of a matrix of individual tasks and is considered verifiable and subject to appeal from employees – the data also confirms that the performance score is a significant function of visible effort – meaning that supervisors face costs from biasing reports on this margin. The attitude score, on the other hand, enters as a

[86] MacLeod and Malcomson [1989]; Prendergast and Topel [1993]; Baker et al. [1994]

[87] Through aggrievance at 'unfair' evaluations, turnover and human-capital loss [Prendergast and Topel, 1993]; through increased noise in performance measures, leading to imperfect monitoring and increased shirking [Prendergast and Topel, 1996]; through increased risk transferred to individuals through the noisier performance measures and a requirement for additional risk compensation [Prendergast and Topel, 1996]; through rent-seeking activities towards impressing the supervisor in non-productive ways [Prendergast and Topel, 1993]; and through worker-task sorting and matching inefficiency as a result of over- and under-evaluated talent [Prendergast, 2002; Frederiksen et al., 2017]

[88] *"We believe there is a great need for more empirical research on the use of implicit contracts and subjective performance evaluation in employment relationships"* [p.1778, Oyer, Schaefer et al., 2011, *The Handbook of Labor Economics*]

single-line input in the appraisal and is unverifiable, meaning that this score is more subject to supervisor bias, at little cost. The data shows substantial heterogeneity in the attitude scores of employees, holding the performance scores fixed, and shows that the employee's position in the distribution of the attitude score relative to her position in the distribution of the performance score responds to relationship characteristics, such as joint tenure and gender, and responds to incentives and monitoring as theorised in the literature. Converting the scores into unitless measures, using z-scores, is necessary since they contain different weights in the overall evaluation. The absolute difference between the attitude z-score and the performance z-score is therefore used as a lower-bound measure of supervisor bias – lower-bound, since it may be possible for the supervisor to adjust the performance score along some dimensions in the matrix, and rational models suggest that this will be in the same direction as the attitude score.

I show that the measure of bias responds to joint tenure and gender; and that this relationship is stronger when there are greater incentives in the organisation (the supervisor's appraisal has a greater impact on the subordinate's monetary outcomes); and weaker when there is more monitoring in place, following Prendergast and Topel [1996] and Prendergast [2002]. A one-standard deviation 'over-valuation' in the appraisal score is associated with a 0.2 increase in the grade level of the civil servant, which equates to approximately a 3% increase in wages relative to the median; a one-standard-deviation increase in the autonomy granted to the management further increases this effect by 0.3 grades or a further 3% increase in the wage relative to the median. These findings are conditional on controlling for the performance level of the employee and are robust to sector and jurisdiction fixed effects, individual controls, management practices as per the World Management Survey, characteristics of the management, and wider clustering of the standard errors.[89] I show that the overvaluation is not associated with selective tenure in the organisation and does not seem to be measuring the intrinsic motivation or the skill level of the employee.

At the organisation level, the results show a significant negative relationship between the organisation-level measure of bias and service delivery outcomes across a range of policy domains:[90] a one unit increase in bias at the organisation level is associated with a 0.04 standard-deviation decrease in service delivery performance.[91] To put this into perspective in terms of service delivery outcomes, this is equivalent to a worsening in the pupil-class ratio by around 1 student or a decrease in the antenatal care rate by 1.2 percentage points (see Table A1). To put this into perspective in terms of alternative

[89]The methodology used for the World Manager Survey follows Bloom and Van Reenen [2007*b*, 2010]; Rasul and Rogger [2016]

[90]The organisation-level measure of bias is the average across the individual-level measures of bias

[91]This means that individuals are 'overvalued' or 'undervalued' relative to the objective performance score by one standard deviation on average

policies, an unconditional regression of the service delivery outcome on aggregate management practices in the same sample shows that a one-standard-deviation increase in aggregate management practices is associated with a 0.08 standard-deviation increase in service delivery outcomes across all policy domains, meaning that bias can cost up to around half of the gains from management practices. The findings are robust to indicator fixed effects, controls for employee characteristics, management practices, management characteristics, jurisdiction characteristics, manager performance scores, uncertainty in the jurisdiction as measured by rainfall variation, and clustering of the standard errors at the sector level.

To further understand under which conditions bias has a greater impact on service delivery outcomes, regressions with interactions show that bias has a larger impact on service delivery when the service delivery environment is more diverse – when ethnic fractionalisation and linguistic fractionalisation are greater, and when there is greater deviation in period-to-period outcomes across service delivery measures in general (uncertainty in the environment) – and when monitoring is less efficient in the organisation – when there are fewer principals to oversee the evaluation process and when the monitoring technology is inefficient (or non-existent). These are conditions under which the productivity of individual civil servants is harder to observe and the probability of the supervisor being punished for biasing the report lower. This provides evidence of the potential benefits to a larger number of principals – in relation to Bernheim and Whinston [1986]; Dixit [2002*a*] – and the potential costs of greater autonomy to management – in relation to Rose-Ackerman [1986] and Rasul and Rogger [2016], when subjective performance evaluations are present.

This paper contributes to two main strands of literature: that on subjective performance evaluations by providing the first individual-level empirical evidence directly related to the theoretical framework put forward in Prendergast and Topel [1996] and Prendergast [2002]; and, more broadly, that on optimal employment contracts, by empirically showing the importance of considering subjective performance evaluations and supervisor when deciding optimal levels of manager autonomy [Rose-Ackerman, 1986; Melumad et al., 1992; Aghion and Tiróle, 1997; Laffont and Martimort, 1998; Mookherjee, 2006; Rasul and Rogger, 2016] and determining optimal contracts under multiple principals with imperfect monitoring [Bernheim and Whinston, 1986; Martimort, 1996; Dixit, 1997].

This paper also contributes to the wider literature on the productivity public officials. While there have been significant strides in the understanding of the effects of monetary and non-financial incentives in terms of the effort response and selection of public-sector workers [Georgellis et al., 2010; Dal Bó et al., 2013; Ashraf, Bandiera and Lee, 2014; Ashraf, Bandiera and Jack, 2014; Finan et al., 2017; Ashraf et al., 2016;

Deserranno, 2017], there is a lot to learn regarding the contracting environment of bureaucrats, particularly in the middle tier [Rasul and Rogger, 2016].

This paper is organised in the following manner: section 5.2 outlines the context of the Ethiopian bureaucracy and details the evaluation system; section 5.3 details the empirical specifications and the findings for the individual-level analysis (5.3.1) and for the organisation-level analysis (5.3.2); section 5.4 provides a discussion of the results.

5.2 Context

5.2.1 Ethiopia

The context of this study is the Ethiopian public sector. Ethiopia is a highly populated country situated n East Africa, with a population of approximately 100 million individuals and a GDP per capita of around 700 USD per person in nominal terms or 1,700 USD in purchasing-power parity terms.

Governance in Ethiopia is centred around three major tiers of government: federal, regional, and woreda (district). Some of the regions have zones, which act as intermediaries between the regional- and woreda-level governments; kebeles are also considered part of the government structure and can be thought of as neighbourhood collectives that interact with both the woreda governments and the communities [Federal Democratic Republic of Ethiopia, 2012]. Since a major drive to decentralise budgets, decision making, and revenue collection through the *District Level Decentralization Program*, the majority of the planning and design of non-federal economic policies is conducted at the woreda level, with woredas considered the socio-economic centres of policy [Ministry of Finance and Economic Development, Ethiopia, 2002; Gebre-Egziabher, 2007; Khan et al., 2014].

Regions are transferred close to half of all public resources and, of this, woredas manage around 80% [World Bank, 2010*a*, 2015*d*]. Almost all of the non-federal expenditure, therefore, flows through the woredas, which have autonomy in its allocation. This makes the woreda-level government the major tier when it comes to non-federal policy.

Table A1 shows the summary statistics from a broad range of demographic and service delivery statistics. The service delivery statistics are used as outcome variables in the organisation-level analysis.

5.2.2 The officials and data coverage

The structure of a public-sector office in Ethiopia is largely uniform across organisations, as guided by the Ministry of Public Service and Human Resource Development.

The head of the organisation is politically appointed and is responsible for the overall performance of the organisation; the head answers to the respective administration or council if the organisation is underperforming.[92] The head of the organisation decides the structure of the incentives in the organisation and the extent to which to delegate to the director. The supervisors are professional civil servants, recruited or promoted to the position of management by the head of the organisation.[93] The head will allocate responsibility to the supervisor and determine the extent to which the supervisor has autonomy over operation decisions, personnel decisions, incentive decisions and so on. The supervisor is responsible for a small team of employees, with an average span of control of 8 professional civil servants (plus administrative staff), see table 1. The recruitment and promotion of professional civil servants are ultimately signed off by the head of the organisation with input from the supervisor, although the extent to which the supervisor has authority over this will vary across organisations. Professional civil servants are recruited by the head and the supervisor to fulfil technical and administrative tasks and report to their supervisor(s). These professional civil servants are annually evaluated by their immediate supervisor.

The study uses data from the Ethiopia Civil Servants Survey (ECSS), which covers a representative sample of civil servants across all three major tiers of government and across five core sectors – agriculture, education, health, revenue, and trade – of service delivery. From this dataset, performance evaluation data was collected from a subset of organisations, which are 'statistically similar' to the rest of the organisations covered in the ECSS (see table 1). Table 1 also shows that the individuals in the evaluation-data sample are largely similar, though statistically significantly different from the rest in the sample in some ways that I control for in all analyses.

5.2.3 Annual evaluations and measuring supervisor bias

The nature of the evaluation system in the Ethiopian bureaucracy allows for the measurement of supervisor bias. The evaluation is conducted by the immediate supervisor for each subordinate individually. The evaluation system takes the form of two individual reports for each employee: an 'attitude' score and a performance score. The performance score is a weighted average across multiple scores which are linked to specific activities and tasks and is relied upon as a verifiable measure of employee performance – see figure A1 for an extract; the performance metric is extremely detailed and often exceeds ten pages (the appendix figure is a two-page extract from a fourteen-

[92]The Prime Minister's Office or National Parliament is the organisation is Federal; The Regional President's Office or Regional Parliament is the office is Regional; and the Woreda administration or Woreda council if the organisation is at the woreda level.

[93]The supervisor is also referred to as the manager, the director, and the process-owner.

page performance appraisal). The attitude score is supposed to measure organisational culture and alignment [Abagisa, 2014] and the "nature and behavior" of the employee [Tereda, 2014] and is entirely unverifiable – the attitude score is a one-line input at the very end of the report, with no further validation or detail.

The dual-report system in Ethiopia allows the measurement of a lower-bound value for the supervisor bias towards an individual employee. Since the performance score is largely verifiable, based on visible metrics, and subject to appeal by the employee if an incorrect or unfair score is provided, I argue that the supervisor bias will largely arise from adjustments to the attitude score. Since the attitude score is entirely unverifiable, the supervisor has little disincentive to adjust the score of favoured or disliked employees along this margin. However, since the scores carry different weights for the final evaluation of the employee – see the distributions of the scores in figure A2 – it is necessary to transform the scores into unitless measures and this is done by creating z-scores for each score for each employee, which represents the employee's position in the distribution of the scores and then taking the absolute difference between the z-scores as the measure of how much an employee is pushed up in the distribution of scores relative to the performance score or pushed down.[94] Further, since it is of interest to understand the differing effects of being pushed up versus pushed down, an indicator for whether the actual difference between the z-score of the attitude score and the z-score of the performance score is positive is also used in the analysis.

An alternative measure of supervisor bias involves using the residuals from a wage regression of the individual civil servant, controlling for the total evaluation score of the civil servant, the tenure of the civil servant in the position, the tenure of the civil servant in the organisation, the tenure of the civil servant in the service, an indicator for whether the civil servant has an undergraduate degree, and indicator for whether the civil servant has a masters degree, and organisation fixed effects, which account for tier, sector, jurisdiction, and organisation-specific characteristics. The remaining variation in the wage of the civil servant should capture unobserved bias if the total evaluation score captures all measures of unobserved productivity. Analysing how these residuals are correlated with the relationship characteristics and the proposed measure of bias, shows that the proposed measure of bias is significantly correlated with this measure and therefore validates the current approach towards measuring supervisor bias. This is shown in table A2.

[94]The attitude score is typically weighted at 40 percent and the performance score at 60 percent, dependent on the sector, which is always controlled for.

5.2.4 Measuring management practices

The measurement of management practices uses the methodology of the World Management Survey, applied to the Ethiopian public-sector context [Bloom and Van Reenen, 2007*b*, 2010; Rasul and Rogger, 2016]. All supervisors in the sample were surveyed the World Management Survey items adapted for this context and Table A3 provides detail on each item.

Three different aggregations of the management practices are used throughout the analysis to understand differing margins of management practices.

Sections 5.3.1.1 and 5.3.1.2 group management practices into incentives and targeting, monitoring and all other management practices (combining flexibility, roles, staff involvement, and staffing) in order to understand how bias responds to management practices where there is more money on the line for the employee as a result of the evaluation, and management practices where there is more efficient monitoring in place to limit supervisor bias.

Section 5.3.1.3 uses the management practices related to roles as a proxy for managerial autonomy (targeting, monitoring, incentives, flexibility, staffing, and staff involvement are included separately) to understand the impact of manager discretion in the organisation on the bias-wage relationship.

All analyses at the organisation-level under section 5.3.2 aggregate management practices across all dimensions in order to control for the efficiency of management practices in the organisation as a whole.

Details of each item and the indices used are provided in Table A3.

5.3 Findings

5.3.1 Individual-level analysis

The analyses below are at the individual level and study the relationship between the scores and individual characteristics; the scores and relationship characteristics between the management and the employee; the response of the relationships to incentives and monitoring; and the impact of the proposed measure of supervisor bias on wages.

Since each sub-section reports the results of a different specification, the specifications are detailed in the sub-sections and the tables.

5.3.1.1 What are the scores correlated with?

This section provides introductory descriptive evidence on how the performance and attitude scores are related to one another and how they vary individually with individual characteristics and management characteristics.

Figure 1 shows that the correlation between the two scores is weak and negative even conditional on organisation fixed effects. This is in line with the preferences of supervisors to compress the wage schedule, as reported in much of the literature on subjective performance evaluations [Prendergast and Topel, 1993, 1996; Bentley MacLeod, 2003; Levin, 2003]. Figure 2 shows the substantial heterogeneity in the scores given to individuals, even for the same performance score, same total score, the same rank of performance score and the same rank of z-score. This heterogeneity illustrates the substantial room for variation along the margin of the attitude score, holding the performance score constant.

Table 2A shows the results from a regression of the z-score of the performance score on individual characteristics (column 1), including an indicator for whether the manager and employee are of the same gender in column 2, the joint tenure in years in the organisation shared by the employee and manager in column 3, the absolute age difference between the employee and manager in column 4, measures of intrinsic motivation in column 5, measures of perceived technical knowledge in column 6, and the attitude score in column 7. Measures of intrinsic motivation include an indicator equal to one if the respondent selects "The chance to serve Ethiopia" as the response to "What most influenced you to take up a career in the service?" and the public service motivation z-score is created using the items from Perry [1996]. Measures of perceived technical knowledge include the response to "Please indicate [on a likert scale from strongly disagree to strongly agree] the extent to which you agree with the following statements: I have sufficient knowledge to effectively perform my tasks"; and an indicator equal to one if the respondent answered "I do not have the right skills to be productive in this role" to "What are the biggest challenges to you being able to complete your most important tasks effectively?". Standard errors are clustered at the organisation level, allowing for the errors to display heteroskedasticity and serial correlation within the organisation, but assuming that they are independent across organisations. Table 2B shows the same for the attitude score as the dependent variable.

The table shows that the performance score is a function of visible effort, while the attitude score does not seem to be significantly and robustly associated with the measure of visible effort. Table 2B also provides indicative evidence that the attitude score is not capturing intrinsic motivation or other soft skills, even showing a significant negative relationship with this measure of intrinsic motivation. Individual characteristics relating to education, tenure, and gender do not appear to be significantly associated with the appraisal scores on either dimension. The final column of Table 2A and 2B again shows the significant negative correlation between the two scores. Finally, looking at the adjusted R-squared for the specifications shows that more of the variation in the performance score is explained by these observable characteristics.

5.3.1.2 Does the extent of bias change with the amount of money on the line or the extent of monitoring?

Table 3 shows the results of a regression of two different dependent variables signifying supervisor bias towards the evaluated employee on individual characteristics, relationship and management practices, also controlling for sector and jurisdiction fixed effects. Individual characteristics include an indicator for whether the respondent has an undergraduate degree; and indicator for whether the respondent has a masters degree; the respondent's tenure in the position the respondent's tenure in the civil service; the gender of the civil servant.

In columns 1 to 3, the dependent variable is an indicator equal to one if the z-score of the attitude score is greater than the z-score of the performance score, representing whether the evaluated employee has been pushed up in the distribution of evaluation scores, relative to the verifiable performance score; in columns 4 to 6, the dependent variable is the absolute difference between the z-score of the attitude score and the z-score of the performance score. Standard errors are clustered at the organisation level.

The z-score is used to ensure a unit-less measure for each score, allowing the relative difference to be a meaningful representation of the extent to which the individual's appraisal score is pushed upwards or downwards by the attitude score, relative to the distribution of the individual within the verifiable performance score. The absolute difference between the two scores is the proposed (lower-bound) measure of supervisor bias.

The results provide significant evidence that this measure of supervisor bias is correlated with gender. The probability of upward bias is significantly lower for females *if they are supervised by a male*, but if they are supervised by a female, the probability of upward bias significantly increases. Female supervisors are also significantly more likely to provide upward biased scores also for males. The divergence between the performance and attitude score based on gender preferences also shows up in columns 5 to 6, where the dependent variable is the absolute difference between the z-scores of the performance and attitude scores.

Figure 3 shows the OLS coefficient on an indicator if both the subordinate and the supervisor are female, relating to column 3, but for each decile of Management practices: Incentives and Targeting. The figure provides empirical evidence that supervisors are more likely to inflate the appraisal score of the subordinate towards their gender preferences if there are greater incentives in place – so that the favouritism is worth more in terms of monetary value – in line with Prendergast and Topel [1993, 1996] and Prendergast [2002]. The management practices of incentives and targeting

and monitoring are based on the World Management Survey, and the formation of the quasi-aggregate indices follow the existing literature [Rasul and Rogger, 2016]; further detail is provided in Table A3.

An alternative measure of supervisor bias, shown in table A2, involves using the residuals from a wage regression of the individual civil servant, controlling for the total evaluation score of the civil servant, the tenure of the civil servant in the position, the tenure of the civil servant in the organisation, the tenure of the civil servant in the service, an indicator for whether the civil servant has an undergraduate degree, and indicator for whether the civil servant has a masters degree, and organisation fixed effects, which account for tier, sector, jurisdiction, and organisation-specific characteristics. The proposed measure of bias is significantly correlated with this wage residual, providing further validation for this measure. The relationship characteristics are not significantly correlated with the wage residual measure. This can be explained by the fact that the supervisor bias affects the wages through the evaluation score, which is partialled out in the wage regression.

5.3.1.3 Does bias affect wages, conditional on performance?

Table 4 takes the proposed measure of supervisor bias as an independent variable in the civil servant's wage regression. Since the wages in the civil service are a straightforward matrix dependent on the grade level of the civil servant, the dependent variable is the grade. The median civil servant in the sample is of grade 6, which represents a wage of 1586 ETB per month or around 60 USD per month. An increase of 1 grade level to grade 7 from the median is equal to a wage increase of 242 ETB per month or around 9 USD, equal to a 15% wage increase.

Table 4 shows the results of a regression of the grade of the civil servant on the overall measure of bias in column 1; and from column 2 onwards shows the positive and negative measure of bias, controlling for the performance score, individual characteristics[95], sector fixed effects, and jurisdiction fixed effects, across all columns; also controlling for management practices related to autonomy (the managerial discretion over roles) and all other management practices (incentives, targeting, monitoring, flexibility, staffing and staff involvement) in column 3 to 6. Standard errors are clustered at the organisation level in columns 1 to 5 and at the jurisdiction level in column 6.

Table 4 shows that the wage is significantly positively associated with the performance score, as expected. In addition, the results provide evidence that the civil ser-

[95]Individual characteristics include an indicator for whether the respondent has an undergraduate degree; and indicator for whether the respondent has a masters degree; the respondent's tenure in the position the respondent's tenure in the civil service; the gender of the civil servant; and the number of hours worked in a typical week by the civil servant.

vant's wage is a positive function of the positive bias of the supervisor, even controlling for actual performance and individual characteristics. Once relationship characteristics are included, this positive association becomes insignificant, highlight the role that relationship characteristics play in determining the level of bias. I then interact the measure of positive bias to study whether greater managerial autonomy affects the extent to which bias can determine wages. The results show that the greater the autonomy provided to management, as proxied by the roles management practices which measure managerial discretion over staff roles, the more they are able to transfer this positive bias into increases wages for the subordinate: a one-standard-deviation increase in the extent of autonomy in the organisation further increases the association between positive bias and wage by 0.2 grade levels (the equivalent of an additional 3% increase in wage). This result is robust to conditioning on individual characteristics, actual performance, management practices, sector fixed effects and jurisdiction fixed effects. This effect is also robust to the clustering of the standard errors at the wider jurisdiction level rather than the organisation level (column 6). These results provide evidence that the supervisor is more able to convert positive bias into wages when she is given greater autonomy, adding an additional set of considerations to the optimal level of discretion to grant management [Rose-Ackerman, 1986; Rasul and Rogger, 2013].

However, as evidenced from the results in column 5, additional autonomy to the manager does not mean that the individual civil servant's performance score is converted into higher wages. The fact that the supervisor does not or is not able to convert the performance score into wages to a greater extent when she is granted greater autonomy provides further credibility that the analysis is capturing the subjective bias of the management towards the subordinates.

Figure 4 shows the coefficient of the bias in the wage regression as per column 4 of Table 4, for each decile of Management practices: roles and the 95% confidence intervals to show graphically the interactive relationship between positive bias and wages through autonomy. The coefficient is only positive and significant at high levels of management practices relating to autonomy, suggesting that it is only when managers are free to act on discretion that bias creeps into wages in meaningful ways.

5.3.1.4 Is the bias instead picking up measures of skills or internal characteristics such as intrinsic motivation?

One concern when analysing the results of Table 4 is whether the measure of bias, measured as the absolute difference between the z-score of the performance score and the z-score of the attitude score, is capturing the extent to which the employee truly has a good attitude and hence maybe more productive through 'softer' skills. Furthermore,

the question of whether selection into longer tenures in the organisation is driving the results is a genuine concern: supervisors may be able to force out certain employees based on preferences and whims and hence the resulting set of employees being endogenously selected into certain management styles and more or less biased managers – that is to say that supervisors may be endogenously driving the selection of workers into the organisation simultaneously affecting bias and wage outcomes.

Table 5 shows the results of a regression of the tenure of the employee in the organisation on positive and negative bias, the performance score, controlling for individual characteristics, sector fixed effects, and jurisdiction fixed effects in column 1.[96] In column 2, the outcome variable is an indicator equal to one if the respondent selects "The chance to serve Ethiopia" as the response to "What most influenced you to take up a career in the service?"'; in column 3, the outcome variable is the public-service motivation score in z-score [Perry, 1996]; in column 4, the response to "Please indicate the extent to which you agree with the following statements: I have sufficient knowledge to effectively perform my tasks" on a scale of 1 to 5; in column 5 an indicator equal to one if the respondent answered "I do not have the right skills to be productive in this role" to "What are the biggest challenges to you being able to complete your most important tasks effectively?"; in column 6 an indicator equal to one if the respondent possesses an undergraduate degree or higher. Standard errors are clustered at the organisation level for all columns.

Table 5 provides evidence against selective tenure (column 1) correlated with the measures of bias; and provides evidence against positive bias measuring higher levels of important 'soft' skills (columns 2 to 6). If anything, these measures of bias are negatively correlated with soft skills.

5.3.1.5 Potential biases due to measurement and omitted variables

Upon observing the performance-score matrix in figure A1, it is valid to be concerned over the extent to which the performance measures themselves can be subjective. I do not rule this out. The measure of bias is considered a lower bound and hence the estimates are considered downward biased. This is because, there is no rational reason why the subjectivity in the performance score would oppose the direction of the subjectivity in the attitude score, and hence they should both be adjusted in the same direction. Hence, the absolute difference would provide an underestimate of the extent to which an employee has been overvalued or undervalued.

[96]Individual characteristics include an indicator for whether the respondent has an undergraduate degree; and indicator for whether the respondent has a masters degree; the respondent's tenure in the position the respondent's tenure in the civil service; the gender of the civil servant; and the number of hours worked in a typical week by the civil servant.

The second set of concerns regard the potential for the attitude score to be picking up something valuable in terms of the employee's potential performance. This is addressed to some extent in the previous sub-section, but is also part of the following analysis which tries to understand whether the extent of bias has a cost on the service delivery performance of the organisation. If the attitude score inflation is picking up something valuable, then one expects the bias measure to have either no explanatory power in the service delivery performance of the organisation or to demonstrate a positive relationship. The following organisation-level analyses address this.

The third set of concerns regard the extent to which bias is endogenously associated with assortative selection within the organisation, between managers and employees, or with selection into the organisation and these are addressed in section 5.3.1.3 and 5.3.1.4 respectively.

The final set of concerns are around the potential endogenous relationship between supervisor bias and supervisor (and therefore employee) productivity in the wage regressions. A valid argument can be that a biased supervisor is also likely to demonstrate differing levels of productivity which determine the potential for the employee to be productive and hence generate a larger wage. The results in Table 4 show evidence to the contrary of this argument, by comparing the results in column 2 and column 3, where the former excludes all manager characteristics and management practices and the latter includes these, one can see that the coefficient of interest on the positive bias remains unchanged and significant; furthermore, the adjusted R-squared remains constant, suggesting that management in general does not seem to be driving the relationship of interest and, more broadly, does not seem to be driving the wage determination of subordinates conditional on the evaluation scores and the other controls.

5.3.2 Organisation-level analysis

This section switches the analysis to the organisation level to understand the potential efficiency costs of supervisor bias. An alternative hypothesis of the supervisor bias is that it represents the manager's valuation of the unobservable skills of the employee; though this is attempted to be addressed under the individual-level analysis, this section provides further evidence against this hypothesis by showing that this measure is costly in terms of service delivery performance – if the measure was a legitimate valuation of some other skills of the employees, this would not be the case.

The unit of observation is the organisation-indicator pair, where the indicator is the z-score of a specific service delivery measure for the organisation relative to all other organisations. Since the service delivery indicators cover distinct areas of policy, the unit of observation is at the indicator level to analyse whether bias is associated

with increases or decreases in terms of service delivery indicators across the range of indicators. To control for unobserved differences across indicators, for example the general difficulty (differences in the production functions) in improving one indicator relative to another, indicator fixed effects are accounted for in the analysis. Since service delivery is determined and measured at the district level and not at the higher tiers of government, this analysis is restricted to district level. The list of service delivery measures and descriptive statistics are provided in Table A1. The indicators used are those collected directly by the government monitoring systems themselves and are, hence, those on which government judges its own performance.

The analyses below regress the z-score of the service delivery performance indicator across all policy domains on the organisation-average level of bias (the organisation average of the absolute difference between the performance z-score and the attitude z-score); the organisation average performance z-score; aggregate management practices across all sub-indices of the World Management Survey to control for difference in aggregate management practices across organisations [Bloom and Van Reenen, 2010, 2007*b*]; organisation averages of employee characteristics; management characteristics; jurisdiction characteristics controlling for population (size), the percentage of rural inhabitants, remoteness and road density, ethnic fractionalisation and the uncertainty of the service delivery environment.

All policy domains are stacked to allow regressions at the individual indicator level rather than averaging across very different service delivery measures. The qualitative results are almost identical when one takes averages across all service delivery indicators and runs an organisation-only-level analysis and are shown in Table A5. Ethnic fractionalisation measures the probability that two randomly selected individuals in the jurisdiction are from different ethnicities (Fearon, 2003). Individual characteristics include the organisation-average values for: the education level of the employees; an indicator for whether the employee is female; the average effort exerted by employees; the tenure in the position; and the tenure in the civil service. Effort is measured by Hours worked during a typical week is the respondent's response to "What is the actual number of hours you work in the civil service in a typical week?". Management characteristics include organisation-level averages of the tenure in the organisation of the management, tenure in the civil service of management, the education level of the management, the number of principals, and the extent of multitasking experienced by employees on average. Jurisdiction characteristics include the population of the jurisdiction; the percentage of rural inhabitants; the ethnic fractionalisation (in z score); the remoteness measured in terms of the average travel time to the nearest urban centre of at least 50,000 people (a census measure); the road density in the jurisdiction; and the organisation-average difference in service delivery between the most recent two

periods of data (averaged across all indicators, for a measure of overall uncertainty). I also control for indicator fixed effects, controlling for all variation ascribed to the specific indicator used as an outcome. Indicator fixed effects also absorb sector effects, since the indicator also defines the relevant sector. Standard errors are clustered at the organisation level.

The qualitative results are almost identical when one takes averages across all service delivery indicators and runs an organisation-only-level analysis – this is shown in table A4.

5.3.2.1 Is bias at the organisation level costly for service delivery performance?

Table 6 shows that there is a significant negative relationship between supervisor bias and service delivery across all specifications and under different clustering of standard errors. The results suggest that a one-standard-deviation-adjustment in the evaluation by the supervisor relative to the verifiable performance score is associated with a 0.04 standard-deviation reduction in the service delivery performance of the organisation. A one-standard-deviation-adjustment is a one-unit increase in the organisation-average absolute difference between the z-score of the performance score and the z-score of the attitude score.

To put this into perspective in terms of service delivery outcomes, this is equivalent to a worsening in the pupil-class ratio by around 1 student or a decrease in the antenatal care rate by 1.2 percentage points (see table A1). To put this into perspective in terms of alternative policies, an unconditional regression of the service delivery outcome on aggregate management practices in the same sample shows that a one-standard-deviation increase in aggregate management practices is associated with a 0.08 standard-deviation increase in service delivery outcomes across all policy domains, meaning that bias can cost up to around *half of the gains from management practices*.

Column 6 provides a further robustness check against uncertainty in the local environment driving the results, controlling for the standard deviation in the rainfall across the year. Note that the sample is smaller for this regression due to missing rainfall data on some jurisdictions. Column 7 provides a further robustness check against manager productivity simultaneously driving the level of bias and the service delivery performance of the organisation by controlling for the performance scores of the management as evaluated by the head of organisation. Note again that this can only be done for a subset of organisations for which there is also management evaluation scores. These robustness checks and identification concerns are discussed further in section 5.3.2.4.

Table A4 shows that the results are qualitative unchanged when conducting this

analysis at the organisation level, using the organisation-average service delivery performance measure, by averaging across all of the z-scores of service delivery performance.

Another striking observation from table 6 is the significant, robust, negative correlation between the organisation-average value of the performance z-score and the service delivery indicator. The fact that an organisation has higher performance scores across the entire distribution of individuals within the organisation may signify leniency. This may impede the effectiveness of incentives in the workplace, though the study of the mechanisms behind this result is outside of the scope of this paper.

5.3.2.2 When is bias more costly and when can the effects of bias be limited?

To further understand where supervisor bias is more of a problem and to further understand potential policy tools for minimising the cost of bias, table 7 interacts the organisation-level bias with proposed mediating factors.

Columns 2 and 3 show that the problem of organisation bias is worse in areas with more diverse preferences (and therefore service delivery demands) as measured by the ethnic and linguistic fractionalisation of the population in the jurisdiction. This follows the conceptual framework of Prendergast [2002] where the bias leads to a misallocation of workers to tasks and hence organisational efficiency costs post-allocation. The results suggest that this mechanism may be at work, since this misallocation of workers to tasks is likely to feed through to worse service delivery outcomes when the service delivery demands are more complex and diverse.

Columns 4 and 5 show that alternative monitoring of worker performance and behaviour may limit the potential costs of supervisor bias. Column 4 shows that a greater number of principals reduces the cost of the bias on service delivery. Column 5 shows that an increase in the *inefficiency* of the monitoring technology increases the cost of bias on service delivery performance. Inefficiency in the monitoring technology is measured by days taken to receive data after a request to the MIS team. These findings suggest that the misallocation of workers to tasks can be limited if there are other reports regarding workers' talents and performances from other principals or through service delivery performance data, implying the potential for the use of 360-peer reviews and monitoring technologies in performance-appraisal systems.

Figure 5 shows these relationships graphically by showing the coefficient of the organisation-level of bias on the service delivery indicator z-score for each decile of the mediating factors. The regressions are equivalent to table 6 column 5. The top-left panel includes an additional control for linguistic fractionalisation and the bottom-right panel includes an additional control for the monitoring technology inefficiency.

Table A5 shows the results for the same analysis, when separating out the positive and negative bias. The results confirm the earlier findings that the inefficiencies caused by supervisor bias appear to operate through positive bias.

5.3.2.3 Is there an association between bias and percentage of budget expended?

An often-used measure of public-sector efficiency is the ability to expend the approved budget. Whether there is any association between the percentage of approved budget expended and organisation-level bias is addressed in table A6.

Table A6 shows the results of a set of regressions, regressing the jurisdiction-level total expenditure as a percentage of jurisdiction-level approved budget on the organisation-level biases to see if there is any association between jurisdictions with a larger intensity of bias within their organisations and the jurisdiction's ability to expend its budget. I use the latest year of budget data available (2011/12). While the coefficient on the measure of organisation bias is negative across all specifications, there does not appear to be a robust relationship between bias and expenditure success.[97]

5.3.2.4 Potential biases due to measurement and omitted variables

The key concerns to the results in the above analyses are: the bias and the service delivery performance may be driven by an unobservable set of characteristics that simultaneously determine measurement of worker performance and service delivery performance, such as (i) uncertainty in the environment and difficulty of service delivery provision or (ii) manager productivity or efficiency.

The results are constant throughout all specifications in table 6, even when including measures for the uncertainty of the environment, such as the organisation-average difference in service delivery between the most recent two periods of data averaged across all indicators, for a measure of overall uncertainty; the ethnic fractionalisation; the population of the district; the percentage of rural inhabitants; and rainfall variability. This suggests that the results are robust to these factors driving the relationship.

A similar argument is applied to (ii), since the coefficient of interest in remains unchanged after including management practices and manager characteristics from column 4 to column 5, and the management performance scores in column 7. For example, a manager who is less attentive to details may revise subordinate evaluations upwards

[97] Although expenditure as a percentage of budget may measure organisational efficacy in some aspects, in the domain of organisational bias it is difficult to understand in which direction the relationship will prevail: on the one hand, a low-bias allocative-efficient organisation may be better able to allocate resources and expend the provided budget; on the other hand, a high-bias organisation may weight the utility of its workers over the utility of its citizens and may prefer to allocate a greater proportion of resources to individual salaries, using the evaluation system as a justification.

and also pay less attention to policy details, generating this negative correlation. When controlling for management practices in the organisation, for example the extent of monitoring or the extent of incentives and targeting, should be correlated to variation such a tendency which should affect the coefficient of interest.

An additional concern might surround selective sorting of managers to employees within the organisation, however, the organisation-level analysis, due to the availability of service delivery outcome data, is restricted to the district level, where there is typically only one manager within the organisation. I look into the issue of selective tenure into the organisation in the individual-level analysis and find no evidence of such a phenomenon.

5.4 Discussion

Most compensation arrangements within organisations depart from standard contracting theory in two ways. Firstly, the evaluation is conducted by a a middle-tier supervisor, who is not the residual claimant of output. Secondly, the evaluation is based on subjective judgements of performance rather than objective measures. This allows supervisor bias to enter into performance evaluations which can create efficiency costs.

This paper contributes to the literature on subjective performance evaluations by providing the first, to my knowledge, direct empirical evidence of supervisor bias and shows how this relates to the core economic theory in this area.

Supervisor bias is shown to be a function of personal and relationship characteristics, in particular gender and is shown to respond to incentives in the organisation. Furthermore, the paper shows that supervisors given greater autonomy are able to convert bias into wages for preferred employees and that this is not a deterministic function of selective tenure.

This paper demonstrates the potential negative consequences of supervisor bias in efficiency losses in terms of worse service delivery outcomes and shows that these losses are greater when preferences are more diverse – hence misallocation of workers to tasks has a greater impact on output – and are smaller when the monitoring technology in the organisation is more efficient – either through increased oversight through more principals and authority in the organisation, or through improvements in the technology.

A one-standard-deviation increase in bias is 'worth' around half a standard-deviation decrease in management in terms of the impact on service delivery, suggesting the existence of large achievable gains on service delivery through the reduction in bias.

Overall, the paper contributes to the theoretical literature with new data on subjective performance evaluations and shows that subjective performance evaluations respond to the incentive environment and the effects of subjective performance evalua-

tions on efficiency are also mediated by the incentive environment. The paper, therefore, highlights the importance of considering subjective performance evaluations when designing optimal contracts.

6 Conclusion and Further Research

This thesis uses new data on public-sector organisations and public-sector workers in Ethiopia to shed further light on the functioning of public-sector organisations in developing countries and the importance of the public sector for labour markets, especially skilled labour markets in developing countries.

Firstly, this thesis provides new evidence on the role of the public sector in developing country labour markets, using the case of Ethiopia. The public sector is a large employer in most economies across the world, particularly for the most-educated groups. This means that public-sector hiring policies will have significant consequences for the wider labour market and private-sector productivity. I provide empirical evidence that attractive public-sector wages favoured toward the most-educated groups can have potentially damaging consequences for private-sector productivity in the short run. Nevertheless, the public sector is often one of the few attractive options for graduates in developing countries and therefore provides the incentives to acquire a higher education in the first place. To achieve the dual objective to incentivise higher levels of education and to disincentivise the most-productive members of the workforce queuing and preparing for attractive public employment rather than working in the private sector, I propose two initial policy recommendations: (i) invest in developing the private sector to generate attractive options for skilled individuals outside of the public sector; and (ii) to minimise the length and burdensome process for entry into public employment to allow skilled individuals to work at the same time as preparing for entry into public employment.

Secondly, this thesis provides the first evidence, to my knowledge, on information acquisition in large organisations. I collect new data and construct novel measures of the information that individual agents have regarding their operating environment. This research highlights the extent of misinformation that public officials have on local conditions and provides suggestive evidence that this feeds into allocation decisions. Combining this data with a field experiment that provides better information to a random subset of civil servants, I provide empirical evidence that information acquisition in the public sector is consistent with the classical theoretical framework in this area. I also evaluate an intervention in the bureaucracy of the education sector to provide further evidence on how better information to bureaucrats affects service delivery over time. This research provides some of the first detailed evidence on how information

to public-service *providers* (rather than to users, which has been studied in Ferraz and Finan [2008], Hastings and Weinstein [2008], and Casey [2015] for example) can affect service delivery outcomes and the constraints to information acquisition in the public sector.

Finally, this thesis uses the evaluation system within the Ethiopian bureaucracy to shed further light on the prevalence of supervisor bias in performance evaluations and the potential effects that this can have on efficiency. I develop a novel individual-level measure of supervisor bias and show that this measure relates to relationship characteristics and the incentive environment, as predicted by the theory in this area. I also show that supervisor bias is associated with efficiency losses, in terms of worse service delivery, at the organisation level. To extend this literature, panel data on individual-level performance evaluations would allow *changes* in performance evaluations (and supervisor bias) to be linked to *changes* in civil servant effort, activities, and performance, providing a deeper understanding of the issues surrounding subjective performance evaluations.

The study of public-sector organisations is inherently linked to the problem of measuring the productivity of public officials. While there have been advances – for example, using project-level data [Rasul and Rogger, 2013; Rasul et al., 2018*a*]; service-delivery data [Ashraf et al., 2016; Gulzar and Pasquale, 2017]; sales figures [Ashraf, Bandiera and Jack, 2014]; attendance [Callen et al., 2015, 2018]; tax collection figures [Khan et al., 2015]; subjective assessments of bureaucrat performance by civil servants, politicians, media professionals, members of business and professional associations, and civil society and think-tanks [Bertrand et al., 2017]; and the performance evaluations of bureaucrats as I do in sections 2 and 5 of this thesis – there still remain advances to be made in measuring the holistic individual-level performance of bureaucrats. With such measures in hand, the questions of 'what is optimal public-sector wage premium?' and 'what is the optimal set of contracts in the public sector?' can be better answered.

Table 60: Characteristics of Civil Servants and Organisations Surveyed

Table 1A: Characteristics of Civil Servants and Organisations

Means and standard deviations in columns (1) and (2); Differences and standard errors from t-test in column (3)

	(1) No Evaluation Data	(2) Evaluation data
Organisational Characteristics		
Number of bureaucrats per organisation	5.62	5.89
	[6.23]	[6.07]
Number of heads per organisation	0.90	0.96
	[0.30]	[0.19]
Number of directors per organisation	1.19	1.27
	[0.98]	[0.95]
Number of employees per organisation	3.53	3.67
	[5.58]	[5.27]
Ratio of employees to heads	3.22	3.34
	[4.64]	[4.27]
Ratio of employees to directors	2.54	2.75
	[1.51]	[0.67]
Span of control (employees per director)	8.18	7.15
	[9.15]	[9.18]
Proportion of employees with the same gender as the manager	0.67	0.72
	[0.35]	[0.30]
Average joint tenure between employees and management	4.50	5.05
	[3.60]	[3.96]
Management practices (Aggregate)	-0.09	-0.04
	[0.59]	[0.68]
Manager tenure in the organisation	8.81	9.17
	[6.63]	[7.66]
Manager tenure in the civil service	15.72	15.78
	[7.61]	[8.42]
Manager education level	4.96	4.99
	[0.43]	[0.38]
Number of organisations	250	132
Official's Characteristics		
Age	35.63	34.03
	[8.75]	[9.13]
Years in position	2.70	2.67
	[2.85]	[2.29]
Years in organisation	7.24	7.71
	[6.99]	[7.62]
Years in civil service	13.39	12.16
	[8.75]	[9.31]
Number of different organisations in service	2.80	2.67
	[1.82]	[2.73]
Grade	6.22	5.69
	[2.11]	[1.70]
Education [post-high-school qualification=1]	1.00	1.00
	[0.06]	[0.05]
Education [undergraduate degree=1]	0.83	0.82
	[0.38]	[0.39]
Education masters degree=1]	0.13	0.05
	[0.33]	[0.23]
Gender [female=1]	0.19	0.24
	[0.40]	[0.43]
Effort	40.82	42.66
	[10.63]	[9.66]
Number of bureaucrats	1447	384

Table 1B: Summary statistics of the Balanced Scorecard scores

Means and standard deviations

	(1) Attitude score	(2) Performance score
Employee	34.60	49.46
	[4.06]	[5.31]
Between organisation standard deviation	3.68	4.81
Within organisation standard deviation	2.39	2.81
Overall standard deviation	4.06	5.31
Number of employees [organisations]	384 [132]	384 [132]

Notes: Table 1A: *** denotes significance at 1%, ** at 5%, and * at 10% level. Columns (1) and (2) show the mean of the variable and the standard deviation in brackets. Column means between column (1) and column (2) and the standard error of the t-statistic in brackets. The unit of observation for the first panel is the civil servant and for the second pane employees and directors are evaluated and hence only these individuals are used for the statistics in the first panel. These statistics are based on the survey sample and not on ad control is taken from the survey question "How many personnel do you manage?"; the total is taken for each organisation if there is more than one director; and the organisation-lev organisation has an equal weight in the `By Organisation' statistics). Heads refers to heads of organisation or deputy heads of organisation. Directors are sometimes acting directo refers to directors and process owners. There are 7 cases where the employee was wrongly interviewed as a head and hence did not face the correct track of the survey, meaning are a manager and thus it is not possible to accurately these 7 cases to manager or employee (but we know that they are not heads); these have been dropped from the analysis s reported by heads of organisation. Post high school qualification refers to having a diploma, technical vocation certificate or any other post-high-school qualification or further educa qualificaion refers to having an undergraduate degree of further level of education. Grade is the official civil service grade of the professional civil servant, ranging from 1 to 17 in the Civil Service Human Resource Statistics Abstract (Ministry of Civil Service) records the overall female employee percentage to be 35%; this statistic includes frontline staff. Conditi high-school education or those only within professional service class, the female employee proportion is recorded as 24% and 27% respectively. Figures are rounded to two decim **Table 1B:** All columns reports means and standard deviaions in brackets. The unit of observation is the individual civil servant. Employee refers to any professional civil servant w individuals. Manager refers to civil servants that are responsible for a team of professional and administrative civil servants, also referred to as Directors, Process Owners, and Coo score format and the overall score is the sum of the attitude score and the performance score.

Table 2A: Predictors of the performance scores of employees. Z-scores.
OLS Estimates
The dependent variable is the z-score of the performance score in all columns
Standard errors are clustered at the organisation level for all columns

	Dependent variable: Performance score					
	(1) Individual characteristics	(2) Same gender	(3) Joint tenure	(4) Age difference	(5) Intrinsic motivation	(6) Technical knowledge
Education [Undergraduate degree=1]	0.069	0.071	0.066	0.073	0.069	0.066
	[0.10]	[0.10]	[0.10]	[0.10]	[0.10]	[0.10]
Education [Masters degree=1]	-0.11	-0.11	-0.11	-0.11	-0.12	-0.11
	[0.18]	[0.18]	[0.18]	[0.18]	[0.17]	[0.17]
Years in position	-0.017	-0.017	-0.016	-0.017	-0.016	-0.019
	[0.019]	[0.019]	[0.019]	[0.019]	[0.019]	[0.019]
Years in civil service	0.0037	0.0038	0.0056	0.0056	0.0056	0.0059
	[0.0057]	[0.0057]	[0.0060]	[0.0059]	[0.0059]	[0.0059]
Gender [Female=1]	-0.13	-0.20	-0.19	-0.19	-0.19	-0.19
	[0.10]	[0.18]	[0.18]	[0.18]	[0.18]	[0.18]
Effort	0.0093	0.0095	0.0099	0.0099	0.0098	0.010
	[0.0051]*	[0.0050]*	[0.0051]*	[0.0052]*	[0.0052]*	[0.0053]*
Same gender		-0.077	-0.078	-0.072	-0.074	-0.042
		[0.17]	[0.17]	[0.17]	[0.17]	[0.17]
Joint tenure			-0.0093	-0.0072	-0.0077	-0.0078
			[0.013]	[0.013]	[0.013]	[0.013]
Age difference				0.0085	0.0083	0.0083
				[0.0055]	[0.0055]	[0.0055]
Intrinsic motivation					0.038	0.047
					[0.10]	[0.10]
Public sector motivation					-0.010	-0.0092
					[0.067]	[0.068]
Task knowledge						-0.088
						[0.055]
Lack of skills						-0.35
						[0.26]
Attitude score						
Jurisdiction fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Sector fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Adjusted R-squared	0.43	0.42	0.42	0.43	0.42	0.42
Observations [clusters]	384 [132]	384 [132]	384 [132]	384 [132]	384 [132]	384 [132]

Notes: *** denotes significance at 1%, ** at 5%, and * at 10% level. All columns show OLS estimates. Standard errors in brackets are clustered at the organisation level for all columns. The unit of observation is the individual civil ser variable is the z score of the performance score in all columns. The analysis is restricted to employees for which there is BSC evaluation data. Effort is measured as hours worked during a typical week and is the respondent's response number of hours you work in the civil service in a typical week?". Individual characteristics include an indicator for whether the respondent has an undergraduate degree; and indicator for whether the respondent has a masters degree; in the position the respondent's tenure in the civil service; the gender of the civil servant; and the number of hours worked in a typical week by the civil servant. Sector fixed effects are indicators for whether the respondent works in an o the agriculture, education, health, revenue, or trade sector. Jurisdiction fixed effects control for the administrative jurisdiction (i.e. the specific district if the respondent is a district-level respondent; the specific region if the respondent is and an indicator for federal if the respondent is a federal respondent). Same gender is an indicator equal to one when the manager and the employee are the same gender; Joint tenure is the minimum amount of time that the manager in the current organisation together. Age difference is the absolute difference in the age of the manager and the employee being evaluated. Intrinsic motivation is an indicator equal to one if the respondent selects "The chance to serve response to "What most influenced you to take up a career in the service?". Public service motivation is in z-scores created using the items from Perry (1996). Task knowledge is the extent to which the respondent feels that they posse knowledge to effectively performs the task as elicited from the following item: "Please indicate the extent to which you agree with the following statements: I have sufficient knowledge to effectively perform my tasks". Lack of skills is an the respondent answered "I do not have the right skills to be productive in this role" to "What are the biggest challenges to you being able to complete your most important tasks effectively?" and zero otherwise; and an indicator equal to possesses an undergraduate degree or higher. Attitude score is the attitude appraisal score of the civil servant in z-scores. Figures are rounded to two significant figures.

Table 2B: Predictors of the attitude scores of employees. Z-scores.
OLS Estimates
The dependent variable is the z-score of the attitude score in all columns
Standard errors are clustered at the organisation level for all columns

	Dependent variable: Attitude score					
	(1) Individual characteristics	(2) Same gender	(3) Joint tenure	(4) Age difference	(5) Intrinsic motivation	(6) Technical knowledge
Education [Undergraduate degree=1]	-0.20	-0.20	-0.19	-0.19	-0.18	-0.18
	[0.12]	[0.12]	[0.12]	[0.12]	[0.12]	[0.12]
Education [Masters degree=1]	0.20	0.20	0.20	0.20	0.22	0.23
	[0.23]	[0.23]	[0.23]	[0.23]	[0.23]	[0.24]
Years in position	0.00063	0.00054	-0.0020	-0.0023	-0.0035	-0.00092
	[0.018]	[0.018]	[0.018]	[0.018]	[0.019]	[0.018]
Years in civil service	-0.000040	-0.000084	-0.0033	-0.0033	-0.0034	-0.0032
	[0.0046]	[0.0047]	[0.0052]	[0.0052]	[0.0051]	[0.0052]
Gender [Female=1]	-0.16	-0.14	-0.14	-0.14	-0.14	-0.13
	[0.15]	[0.22]	[0.23]	[0.23]	[0.23]	[0.23]
Effort	0.0014	0.0013	0.00054	0.00053	0.00059	0.00036
	[0.0057]	[0.0058]	[0.0059]	[0.0059]	[0.0059]	[0.0060]
Same gender		0.034	0.036	0.038	0.041	0.039
		[0.21]	[0.21]	[0.21]	[0.21]	[0.22]
Joint tenure			0.017	0.018	0.019	0.018
			[0.015]	[0.015]	[0.015]	[0.015]
Age difference				0.0030	0.0033	0.0030
				[0.0055]	[0.0054]	[0.0055]
Intrinsic motivation					-0.088	-0.091
					[0.070]	[0.072]
Public sector motivation					0.010	0.0090
					[0.062]	[0.062]
Task knowledge						-0.024
						[0.061]
Lack of skills						0.29
						[0.33]
Performance score						
Jurisdiction fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Sector fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Adjusted R-squared	0.35	0.35	0.35	0.35	0.34	0.34
Observations [clusters]	384 [132]	384 [132]	384 [132]	384 [132]	384 [132]	384 [132]

Notes: *** denotes significance at 1%, ** at 5%, and * at 10% level. All columns show OLS estimates. Standard errors in brackets are clustered at the organisation level for all columns. The unit of observation is the individual civil servant. The dependent variable is the z columns. The analysis is restricted to employees for which there is BSC evaluation data. Effort is measured as hours worked during a typical week and is the respondent's response to "What is the actual number of hours you work in the civil service in a typical week?". Ind an indicator for whether the respondent has an undergraduate degree; and indicator for whether the respondent has a masters degree; the respondent's tenure in the position the respondent's tenure in the civil service; the gender of the civil servant; and the number of hour the civil servant. Sector fixed effects are indicators for whether the respondent works in an organisation that is within the agriculture, education, health, revenue, or trade sector. Jurisdiction fixed effects control for the administrative jurisdiction (i.e. the specific district if the re respondent; the specific region if the respondent is a regional respondent; and an indicator for federal if the respondent is a federal respondent). Same gender is an indicator equal to one when the manager and the employee are the same gender; Joint tenure is the minimu manager and employee have spent in the current organisation together. Age difference is the absolute difference in the age of the manager and the employee being evaluated. Intrinsic motivation is an indicator equal to one if the respondent selects "The chance to serve E "What most influenced you to take up a career in the service?". Public service motivation is in z-scores created using the items from Perry (1996). Task knowledge is the extent to which the respondent feels that they posses the necessary knowledge to effectively performs following item: "Please indicate the extent to which you agree with the following statements: I have sufficient knowledge to effectively perform my tasks". Lack of skills is an indicator equal to one if the respondent answered "I do not have the right skills to be productive in th challenges to you being able to complete your most important tasks effectively?" and zero otherwise; and an indicator equal to one if the respondent possesses an undergraduate degree or higher. Performance score is the attitude appraisal score of the civil servant in z-scc significant figures.

Table 3: Predictors of the probability of `upward bias'

OLS Estimates

The dependent variable is an indicator equal to one if the employee is `upwards biased' in the evaluation and zero if not in columns 1 to 3 and the absolute difference between the z-score of the performance score and the z-score of the attitude score in columns 4 to 6

Standard errors are clustered at the organisation level for all columns

	Dependent variable: 1[attitude z-score > performance z-score]			Dependent variable: absolute difference between attitude z-score and performance z-score		
	(1) Baseline	(2) Same gender if both female	(3) Conditional on management practice	(4) Baseline	(5) Same gender if both female	(6) Conditional on management practice
Effort	-0.0011	-0.00094	-0.0010	0.00035	0.00059	0.00061
	[0.0027]	[0.0027]	[0.0027]	[0.0065]	[0.0064]	[0.0063]
Gender [Female=1]	-0.12	-0.26	-0.25	-0.021	-0.33	-0.31
	[0.096]	[0.11]**	[0.12]**	[0.21]	[0.32]	[0.33]
Same gender	-0.095	-0.22	-0.21	-0.092	-0.37	-0.34
	[0.089]	[0.11]**	[0.11]*	[0.20]	[0.30]	[0.30]
Joint tenure	0.012	0.011	0.014	0.017	0.015	0.019
	[0.0086]	[0.0086]	[0.0090]	[0.017]	[0.016]	[0.015]
Age difference	-0.0023	-0.0016	-0.0018	-0.011	-0.0097	-0.010
	[0.0034]	[0.0035]	[0.0034]	[0.0056]**	[0.0057]*	[0.0059]*
Same gender x Gender [Female=1]		0.51	0.49		1.13	1.09
		[0.22]**	[0.23]**		[0.45]**	[0.46]**
Individual characteristics	Yes	Yes	Yes	Yes	Yes	Yes
Management practices	No	No	Yes	No	No	Yes
Sector fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Jurisdiction fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Adjusted R-squared	0.20	0.21	0.21	0.27	0.29	0.28
Observations [clusters]	384 [132]	384 [132]	384 [132]	384 [132]	384 [132]	384 [132]

Notes: *** denotes significance at 1%, ** at 5%, and * at 10% level. All columns show OLS estimates. Standard errors in brackets are clustered at the organisation level for all columns. The unit of observation is the individual civil servant. The dependent variable is a binary variable equal to one if the employee has a higher attitude z-score than performance z-score, and is therefore deemed to be upward biased in the evaluation, in columns 1 to 3; and the absolute difference between the z-score of the attitude score and the z-score of the performance score in columns 4 to 6. The analysis is restricted to employees for which there is BSC evaluation data. Individual characteristics include an indicator for whether the respondent has an undergraduate degree; and indicator for whether the respondent has a masters degree; the respondent's tenure in the position the respondent's tenure in the civil service; the gender of the civil servant; and the number of hours worked in a typical week by the civil servant. Effort is measured by Hours worked during a typical week is the respondent's response to "What is the actual number of hours you work in the civil service in a typical week?". Management practices: Incentives and Targeting refers to the organisation average score for a quasi-aggregate management index combining incentives and targeting for the World Management Survey indices; Management practices: Monitoring is the organisation average of the monitoring score; and Management practices: Other combines roles, flexibility, staffing, and staff involvement; both are in z-scores (Bloom and Van Reenen, 2007; Bloom et al 2012). Same gender is an indicator equal to one when the manager and the employee are the same gender; joint tenure in the organisation, is the minimum amount of time that the manager and employee have spent in the current organisation together. Sector fixed effects are indicators for whether the respondent works in an organisation that is within the agriculture, education, health, revenue, or trade sector. Jurisdiction fixed effects control for the administrative jurisdiction (i.e. the specific district if the respondent is a district-level respondent; the specific region if the respondent is a regional respondent; and an indicator for federal if the respondent is a federal respondent). Figures are rounded to two significant figures.

Table 4: Wages and supervisor bias, interacted with management autonomy

OLS Estimates

The dependent variable is the grade (wage) of the civil servant

Standard errors are clustered at the organisation level in columns 1 to 5 and at the jurisdicition level in column 6

	(1) Wages and bias	(2) Positive and negative bias	(3) Relationship characteristics	(4) Manager autonomy and bias	(5) Manager autonomy and performance	(6) Clustering at the jurisdiction level
Bias	0.021					
	[0.084]					
Performance score	0.095	0.28	0.29	0.27	0.29	0.27
	[0.090]	[0.11]**	[0.12]**	[0.12]**	[0.12]**	[0.12]**
Positive bias		0.21	0.17	0.15	0.16	0.15
		[0.13]*	[0.14]	[0.14]	[0.14]	[0.15]
Negative bias		-0.090	-0.11	-0.096	-0.11	-0.096
		[0.092]	[0.092]	[0.076]	[0.092]	[0.097]
Management practices: Roles				-0.12	0.027	-0.12
				[0.19]	[0.16]	[0.22]
Interactions between bias and management practices						
Positive bias x Management practices: Roles				0.22		0.22
				[0.12]*		[0.12]*
Negative bias x Management practices: Roles				0.086		0.086
				[0.10]		[0.11]
Performance score x Management practices: Roles					-0.11	
					[0.096]	
Individual characteristics	Yes	Yes	Yes	Yes	Yes	Yes
Sector fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Jurisdiction fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Management practices	No	No	Yes	Yes	Yes	Yes
Relationship characteristics	No	No	No	Yes	Yes	Yes
Adjusted R-squared	0.43	0.43	0.43	0.42	0.42	0.42
Observations [clusters]	384 [132]	384 [132]	384 [132]	384 [132]	384 [132]	384 [39]

Notes: *** denotes significance at 1%, ** at 5%, and * at 10% level. All columns show OLS estimates. Standard errors in brackets are clustered at the organisation level for columns 1 to 5 and clustered at the jurisdiction level in column 6. The unit of observation is the individual civil servant. The dependent variable is the grade (wage) of the individual civil servant in all columns. The analysis is restricted to employees for which there is BSC evaluation data. Positive bias is the interaction between the absolute difference between the z score of the performance score and the z score of the attitude score and an indicator for whether the actual difference is positive (z_attitude-z_performance>=0); Negative bias is similar, with an interaction for whether the actual difference is negative (z_attitude-z_performance<0). Individual characteristics include an indicator for whether the respondent has an undergraduate degree; and indicator for whether the respondent has a masters degree; the respondent's tenure in the position, the respondent's tenure in the civil service; the gender of the civil servant; and the number of hours worked in a typical week by the civil servant. Effort is measured by Hours worked during a typical week is the respondent's response to "What is the actual number of hours you work in the civil service in a typical week?". Management practices: Roles refers to the organisation average score for the roles component of the World Management Survey indices; I also control for other management practices, including the organisation average score for incentives, targeting, monitoring, flexibility, staffing, and staff involvement and all are in z-scores (Bloom and Van Reenen, 2007; Bloom et al 2012). Jurisdiction fixed effects control for the administrative jurisdiction (i.e. the specific district if the respondent is a district-level respondent; the specific region if the respondent is a regional respondent; and an indicator for federal if the respondent is a federal respondent). Relationship characteristics are: the same gender - an indicator equal to one when the manager and the employee are the same gender; joint tenure in the organisation - the minimum amount of time that the manager and employee have spent in the current organisation together; age difference - the absolute difference in the age of the manager and the employee being evaluated. Figures are rounded to two significant figures.

Table 5: Are we just measuring values and culture in a legitimate way (as intended by the attitude
OLS Estimates

The dependent variable is the tenure in the organisation of the civil servant in column (1); the intrinsic motivation in column (2); the public service motivation z-score in column (3); the extent to which the respondent feels they have sufficient knowledge in column (4); indicator for whether the respondent feels a lack of skills to be productive is a bottleneck in column (5); an indicator for whether the respondent possesses an undergraduate degree or higher in column (6)

Standard errors are clustered at the organisation level for

	(1) Tenure in organisation	(2) Intrinsic motivation	(3) Public service motivation	(4) Sufficient knolwedge for tasks	(5) Do not have the right skills	(6) Education [Undergraduate degree or higher=]
Positive bias	0.51	-0.11	0.018	-0.059	0.016	-0.065
	[0.52]	[0.045]**	[0.066]	[0.061]	[0.020]	[0.034]*
Negative bias	0.23	-0.023	0.0092	0.031	-0.0065	0.027
	[0.38]	[0.028]	[0.043]	[0.044]	[0.012]	[0.034]
Performance score	0.066	-0.043	0.0037	-0.11	0.0040	-0.042
	[0.59]	[0.051]	[0.073]	[0.063]*	[0.0099]	[0.037]
Individual characteristics	Yes	Yes	Yes	Yes	Yes	Yes
Sector fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Jurisdiction fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Adjusted R-squared	0.61	0.024	0.11	0.087	-0.033	0.11
Observations [clusters]	384 [132]	384 [132]	384 [132]	384 [132]	384 [132]	384 [132]

Notes: *** denotes significance at 1%, ** at 5%, and * at 10% level. All columns show OLS estimates. Standard errors in brackets are clustered at the organisation level for all columns. The unit of observation is the individual civil servant. Intrinsic motivation is an indicator equal to one if the respondent selects "The chance to serve Ethiopia" as the response to "What most influenced you to take up a career in the service?"; the public service motivation z-score is created using the items from Perry (1996); The extent to which the respondent feels that they posses the necessary knowledge to effectively performs the task is elicited from the following item: "Please indicate the extent to which you agree with the following statements: I have sufficient knowledge to effectively perform my tasks"; The dependent variable in column (5) is an indicator equal to one if the respondent answered "I do not have the right skills to be productive in this role" to "What are the biggest challenges to you being able to complete your most important tasks effectively?" and zero otherwise; the dependent variable in column (6) is an indicator equal to one if the respondent possesses an undergraduate degree or higher. The analysis is restricted to employees for which there is BSC evaluation data. Positive bias is the interaction between the absolute difference between the z score of the performance score and the z score of the attitude score and an indicator for whether the actual difference is positive (z_attitude-z_performance>=0); Negative bias is similar, with an interaction for whether the actual difference is negative (z_attitude-z_performance<0). Individual characteristics include an indicator for whether the respondent has an undergraduate degree; and indicator for whether the respondent has a masters degree; the respondent's tenure in the position, the respondent's tenure in the civil service; the gender of the civil servant; and the number of hours worked in a typical week by the civil servant; the education indicators are not included as controls in column (6) as the education indicator is used as an outcome variable. Effort is measured by Hours worked during a typical week is the respondent's response to "What is the actual number of hours you work in the civil service in a typical week?" Sector fixed effects are indicators for whether the respondent works in an organisation that is within the agriculture, education, health, revenue, or trade sector. Jurisdiction fixed effects control for the administrative jurisdiction (i.e. the specific district if the respondent is a district-level respondent; the specific region if the respondent is a regional respondent; and an indicator for federal if the respondent is a federal respondent). Figures are rounded to two significant figures.

Table 6: Services and supervisor bias at the organisation-indicator level

OLS Estimates

The dependent variable is the z-score of the service delivery indicator across all policy domains

Standard errors clustered at the organisation level in columns (1) to (8) and column (10); Standard errors clustered at the service delivery indicator level in column

	(1) Unconditional	(2) Indicator Fixed effects	(3) Employee characteristics	(4) Woreda characteristics	(5) Management	(6) Standard deviation in rainfall across the year (uncertainty)	(7) Management performance scores	(8) Cluster at the indicator level	(9) Cluster at the sector level	(10) Positive and negative bias
Bias	-0.045	-0.045	-0.040	-0.053	-0.050	-0.052	-0.067	-0.050	-0.050	
	[0.026]*	[0.027]*	[0.027]	[0.026]**	[0.027]*	[0.028]*	[0.037]*	[0.027]*	[0.021]*	
Average performance score	-0.16	-0.16	-0.16	-0.14	-0.14	-0.15	-0.060	-0.14	-0.14	-0.18
	[0.030]***	[0.031]***	[0.033]***	[0.031]***	[0.034]***	[0.035]***	[0.085]	[0.057]**	[0.017]***	[0.042]***
Positive bias										-0.073
										[0.031]**
Negative bias										-0.0011
										[0.043]
Indicator fixed effects	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Employee characteristics	No	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Jurisdiction characteristics	No	No	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Management practices	No	No	No	No	Yes	Yes	Yes	Yes	Yes	Yes
Manager characteristics	No	No	No	No	Yes	Yes	Yes	Yes	Yes	Yes
Standard deviation in rainfall	No	No	No	No	No	Yes	No	No	No	No
Manager performance score	No	No	No	No	No	No	Yes	No	No	No
Adjusted R-squared	0.016	0.084	0.086	0.10	0.10	0.11	0.100	0.10	0.10	0.10
Observations [clusters]	1650 [118]	1650 [118]	1650 [118]	1650 [118]	1650 [118]	1594 [114]	1338 [96]	1650 [15]	1650 [5]	1650 [118]

Notes: *** denotes significance at 1%, ** at 5%, and * at 10% level. All columns show OLS estimates. Standard errors are clustered at the organisation level in columns (1) to (8) and column (10); Standard errors clustered at the service delivery indicator level in column (8) and at the sector level in column (9). The unit of observation is at the organisation-indicator level. The dependent variable is the z score of the service-delivery indicator, where the z score is defined within the indicator and across jurisdictions. The analysis is restricted to organisations for which there is BSC evaluation data. Bias is the absolute difference between the z score of the performance score and the z score of the attitude score. Positive bias is the interaction between the absolute difference between the z score of the performance score and the z score of the attitude score and an indicator for whether the actual difference is positive (z_attitude-z_performance>=0); Negative bias is similar, with an interaction for whether the actual difference is negative (z_attitude-z_performance<0). Individual characteristics include the organisation-average values for: the education level of the employees; an indicator for whether the employee is female; the average effort exerted by employees; the tenure in the position; and the tenure in the civil service. Effort is measured by Hours worked during a typical week is the respondent's response to "What is the actual number of hours you work in the civil service in a typical week?". Management practices refer to the aggregate management practices index based on the World Management Survey (Bloom and Van Reenen, 2007; Bloom et al 2012). Management characteristics include organisation-level averages of the tenure in the organisation of the management, tenure in the civil service of management, the education level of the management, the number of principals, and the extent of multitasking experienced by employeed on average. Sector fixed effects are indicators for whether the respondent works in an organisation that is within the agriculture, education, health, revenue, or trade sector. Jurisdiction characteristics include the population of the jurisdiction; the percentage of rural inhabitants; the ethnic fractionalisation (in z score); the remoteness measured in terms of the average travel time to the nearest urban centre of at least 50,000 people (a census measure); the road density in the jurisdiction; and the organisation-average difference in service delivery between the most recent two periods of data (averaged across all indicators, for a measure of overall uncertainty). Ethnic fractionalisation measures the probability that two randomly selected individuals in the jurisdiction are from different ethnicities (Fearon, 2003). The standard deviation in rainfall across the year uses the full set of observations for 2015 from the World Health Organisation LEAP programme and calculates the standard deviation. For a subset of organisations, we also have the performance scores of the managers, as evaluated by the head of organisation, included as a control in column (7). Figures are rounded to two significant figures.

Table 7: Services and supervisor bias. Conditions in which bias is more or less costly

OLS Estimates

The dependent variable is the z-score of the service delivery indicator across all policy domains

Standard errors clustered at the organisation level in all columns

	(1) Baseline	(2) Ethnic fractionalisation	(3) Linguistic fractionalisation	(4) Number of principals (monitoring)	(5) Monitoring technology
Bias	-0.050	-0.10	-0.097	-0.20	0.021
	[0.027]*	[0.025]***	[0.026]***	[0.055]***	[0.033]
Average performance score	-0.14	-0.16	-0.16	-0.15	-0.14
	[0.034]***	[0.032]***	[0.033]***	[0.032]***	[0.030]***
Ethnic fractionalisation	0.028	0.16	0.24	0.028	0.042
	[0.043]	[0.047]***	[0.18]	[0.042]	[0.040]
Bias x Ethnic fractionalisation		-0.11			
		[0.026]***			
Linguistic fractionalisation			-0.084		
			[0.17]		
Bias x Linguistic fractionalisation			-0.098		
			[0.026]***		
Number of principals				-0.076	
				[0.051]	
Bias x Number of principals				0.066	
				[0.022]***	
Inefficiency of monitoring technology					0.00027
					[0.0011]
Bias x Inefficiency of monitoring technology					-0.0015
					[0.00058]**
Indicator fixed effects	Yes	Yes	Yes	Yes	Yes
Employee characteristics	Yes	Yes	Yes	Yes	Yes
Jurisdiction characteristics	Yes	Yes	Yes	Yes	Yes
Management practices	Yes	Yes	Yes	Yes	Yes
Manager characteristics	Yes	Yes	Yes	Yes	Yes
Standard deviation in rainfall	No	No	No	No	No
Manager performance score	No	No	No	No	No
Adjusted R-squared	0.10	0.10	0.10	0.10	0.11
Observations	1650 [118]	1650 [118]	1650 [118]	1650 [118]	1635 [117]

Notes: *** denotes significance at 1%, ** at 5%, and * at 10% level. All columns show OLS estimates. Standard errors clustered at the organisation level in all columns. The unit of observation is at the organisation-indicator level. The dependent variable is the z score of the service-delivery indicator, where the z score is defined within the indicator and across jurisdictions. The analysis is restricted to organisations for which there is BSC evaluation data. Bias is the absolute difference between the z score of the performance score and the z score of the attitude score. Positive bias is the interaction between the absolute difference between the z score of the performance score and the z score of the attitude score and an indicator for whether the actual difference is positive (z_attitude-z_performance>=0); Negative bias is similar, with an interaction for whether the actual difference is negative (z_attitude-z_performance<0). Individual characteristics include the organisation-average values for: the education level of the employees; an indicator for whether the employee is female; the average effort exerted by employees; the tenure in the position; and the tenure in the civil service. Effort is measured by Hours worked during a typical week is the respondent's response to "What is the actual number of hours you work in the civil service in a typical week?". Management practices refer to the aggregate management practices index based on the World Management Survey (Bloom and Van Reenen, 2007; Bloom et al 2012). Management characteristics include organisation-level averages of the tenure in the organisation of the management, tenure in the civil service of management, the education level of the management, the number of principals, and the extent of multitasking experienced by employeed on average. Sector fixed effects are indicators for whether the respondent works in an organisation that is within the agriculture, education, health, revenue, or trade sector. Jurisdiction characteristics include the population of the jurisdiction; the percentage of rural inhabitants; the ethnic fractionalisation (in z score); the remoteness measured in terms of the average travel time to the nearest urban centre of at least 50,000 people (a census measure); the road density in the jurisdiction; and the organisation-average difference in service delivery between the most recent two periods of data (averaged across all indicators, for a measure of overall uncertainty). Ethnic fractionalisation measures the probability that two randomly selected individuals in the jurisdiction are from different ethnicities (Fearon, 2003). Number of principals is the organisation average response to "How many people would you say regularly give you tasks as part of your formal work duties?". Inefficiency of monitoring technology is the organisation average response to "If you needed to request statistical data from the MIS (Management Information System) Directorate, how long do you expect it would take to get (assuming the data exists)?". Figures are rounded to two significant figures.

Table A1: Description of Survey and Administrative Data

Summary statistics

	(1) Mean	(2) Standard deviation	(3) Skewness	(4) Autogregressive parameter from DF regression, no time trends	(5) Autogregressive parameter from DF regression, linear time trends	(6) Z Statistic from Harris Tsavalis Unit Root Test (linear time trends)	(7) P-value from Harris Tsavalis Unit Root Test (linear time trends)	(8) Fixed-N exact critical value for the Im-Pesaran-Shin test at the 5% level (linear time trends)
Demographic and Socio-economic Characteristics								
Population	99605	61868	1.19					
Proportion of rural inhabitants	0.84	0.26	-2.66					
Formal unemployment rate (percentage)	4.07	5.63	2.21					
Education indicators								
Primary enrolment	21928.71	14019.79	1.06	0.61	0.05	-12.36	0.00	-2.34
Primary pupil-class ratio	57.17	26.71	5.32	0.41	-0.03			
Primary pupil-per-school rate	549.92	228.75	1.61	0.57	0.16	-5.46	0.00	-2.34
Primary pupil-teacher ratio	97.61	353.89	15.1	-0.15	0.02			
Health indicators								
Antenatal care rate	0.51	0.29	0.08	-0.13	-0.53	-6.81	0.00	
Contraceptive acceptance rate	0.59	0.29	-0.23	-0.19	-0.52	-6.38	0.00	
Rate of births delivered by skilled attendant	0.33	0.22	2.62	-0.19	-0.60	-8.26	0.00	
Proportion of infants fully immunized	0.74	0.22	-1.24	-0.08	-0.57	-8.06	0.00	
Agriculture indicators								
Hectares of land used for agricultural purposes	29366	33892	6.49	1.01	1.01			
Hectares of land used for pastoral purposes	27009	75005	5.97	1.00	1.00			
Agricultural income per year per household in Birr	11738	15868	4.58	0.32	0.32			
Percent of households in subsistence agriculture	24.29	23.23	1.56	0.92	0.92			
Revenue indicators								
Number of Tax Identification Numbers issued in last financial year	425.45	477.85	2.42	0.33	0.33			
Proportion of total actual income from regional recurrent block grant	0.81	0.18	-1.97	-0.22	-0.22			
Proportion of total actual income from own sources of revenue	0.20	0.15	2.46	-0.08	-0.08			
Trade indicators								
Number of business licenses issued in last financial year	375.82	480.71	5.19	1.36	1.36			
Total revenue from issuing and renewing business licenses last financial year	108473	218841	8.69	0.32	0.32			

Notes: Unit of observation is the distict. Skewness refers to Pearson's moment coefficient of skewness. Column 4 reports the autogregressive coefficient from the standard Dickey Fuller regression in a panel / time-series dataset, controlling for district-specific fixed effects. Column 5 reports the autogregressive regression in a panel / time-series dataset, controlling for district-specific fixed effects and district-specific linear time trends. Column 6 reports the test statistics from the Harris Tsavaliz test, which tests for the existence of a unit-root in a panel dataset, allowing for district-specific fixed effects and linear district-spe common autogregressive parameter for each series; assumes that the number of districts tends to infinity while the total number of time periods is fixed and the errors are serially uncorrelated; the test can only be conducted on a balanced panel dataset; the null hypothesis is that there exists a (common) unit root Harris Tsavalis test, corresponding to the Z-Statistic reports in Column 5. Column 8 reports the fixed-N critical value for the Im_Pesaran-Shin Test of a unit root and Column 9 reports the test statistic for the series, allowing for district-specific fixed effects and district-specific linear time trends, and a district-specif Pesaran-Shin test allows for for panel-specific autoregressive parameters and an unblanaced panel dataset; the Dickey Fuller regression is fitted to each panel separately and an average test statistic is used, under the assumption that the errors are serially uncorrelated, with fixed N and fixed T; the null hypothes root, against the alternative that some of these series contain unit roots. Test statistics for the pupil-class ratio and pupil-teacher ratio are not displayed due to gaps in the panel series; test statistics and autoregressive parameters are not shows for the demographic indicators as these are from a single observatio and autoregressive parameters are not shows for the agriculture, revenue, or trade indicators as these are from a a two-period panel dataset, with insufficent observations to conduct the tests. Figures are rounded to two decimal places.

Table A2: Bias as the wage residual and the proposed measure of bias

OLS Estimates

The dependent variable is the residual of the wage of the civil servant after controlling for the tenure, education, effort, and actual evaluation scores of the individual employee, and organisation fixed effects

Standard errors are clustered at the organisation level for all columns

	(1) Relationship characteristics	(2) Absolute difference in z-scores
Joint tenure in the organisation (Manager and employee)	0.0063	0.0060
	[0.0056]	[0.0057]
Age difference	-0.0087	-0.0083
	[0.0050]*	[0.0051]
Gender [Female=1]	0.016	0.012
	[0.094]	[0.095]
Same gender (Manager and employee)	0.028	0.032
	[0.038]	[0.042]
Same gender x Gender [Female=1]	-0.046	-0.071
	[0.090]	[0.10]
Bias		0.061
		[0.026]**
Management practices	No	No
Adjusted R-squared	-0.0075	-0.0067
Observations [clusters]	383 [132]	383 [132]

Notes: *** denotes significance at 1%, ** at 5%, and * at 10% level. All columns show OLS estimates. Standard errors in brackets are clustered at the organisation level for all columns. The unit of observation is the individual civil servant. The dependent variable is the residual of the wage, after having controlled for the tenure in the position of the employee, the tenure in the organisation, the tenure in the civil service, whether the employee has an undergraduate degree, whether the employee has a masters degree, the effort of the employee, the actual performance score of the employee, and the actual attitude score of the employee, controlling for organisation fixed effects, which control for all management practices and other organisational characteristics. The analysis is restricted to employees for which there is BSC evaluation data. Effort is measured by Hours worked during a typical week is the respondent's response to "What is the actual number of hours you work in the civil service in a typical week?". Same gender is an indicator equal to one when the manager and the employee are the same gender; joint tenure in the organisation, is the minimum amount of time that the manager and employee have spent in the current organisation together; age difference is the absolute difference in the age of the management and age of the employee. Organisation fixed effects control for the tier, sector, and jurisdiction of the organisation, as well as all other organisation-level effects, such as the characteristics of the heads of organisation and management. Figures are rounded to two significant figures.

Table 70: World Management Survey Items and Indices

Table A3: World Management Survey Indicators

Quasi-aggregate index Incentives Monitoring and Targeting	Topic	Item	Comparable to a question used in BVR?
Management practices: Incentives and targeting	Targeting	Does your Directorate have a clear set of targets derived from the organization's goals and objectives? Are they used to determine your work schedule?	Yes
Management practices: Incentives and targeting	Targeting	When you arrive at work each day, do you and your colleagues know what their individual roles and responsibilities are in achieving the organisation's goals?	Yes
Management practices: Incentives and targeting	Targeting	How are targets and performance measures communicated to staff in your directorate?	Yes
Management practices: Monitoring	Monitoring	In what kind of ways does your Directorate track how well it is delivering services? Can you give me an example?	Yes
Management practices: Monitoring	Monitoring	Are you involved in performance review for your Directorate? If so, how often does this occur?	Yes
Management practices: Incentives and targeting	Performance incentives	How would under-performance be tolerated in your Directorate? Can you give me an example of how such a case would be dealt with?	Yes
Management practices: Incentives and targeting	Performance incentives	Given past experience, have members of *[respondent's organisation]* been disciplined for breaking the rules of the civil service?	
Management practices: Incentives and targeting	Performance incentives	Does your Directorate use performance, targets, or indicators for tracking and rewarding (financially or non-financially) the performance of its employees?	
Management practices: Other	Roles	When staff in your Directorate are given tasks in their daily work, how much discretion do they have to carry out their assignments? Can you give me an example?	Yes
Management practices: Other	Roles	Can most staff in your Directorate make substantive contributions to the policy formulation and implementation process?	Yes
Management practices: Other	Roles	Is the workload of achieving your Directorate's targets evenly distributed across its different employees, or do some groups consistently shoulder a greater burden than others?	Yes
Management practices: Other	Roles	Thinking about all the projects that your Directorate has been involved in since your appointment here, would you say that managers and supervisors try to use the right staff for the right job?	Yes
Management practices: Other	Flexibility	Does your Directorate make efforts to adjust to the specific needs and specific requirements of communities, clients, or other stakeholders?	Yes
Management practices: Other	Flexibility	How flexible would you say your Directorate is in terms of responding to new and improved work practices or reforms?	Yes
Management practices: Other	Staff involvement / Contribution	How do problems in your directorate get exposed and fixed?	Yes
Management practices: Other	Staff involvement / Contribution	What kind of feedback do you get in staff meetings?	Yes
Management practices: Other	Staff involvement / Contribution	Let's say you've agreed to a follow up plan at one of your meetings, what would happen if the plan wasn't enacted?	
Management practices: Other	Staffing	Do you think the management of your Directorate think about attracting talented people to your Directorate and then doing their best to keep them? For example, by ensuring they are happy and engaged with their work.	Yes
Management practices: Other	Staffing	If two senior level staff joined your Directorate five years ago and one was much better at their work than the other, would he/she be promoted through the service faster?	Yes

Notes: * This question is not score on a scale from 1 to 5, but is simply a coded response where the least frequent possible response is "Not involved in performance review" and the most frequent possible response is "Weekly". World Management Indicators based on Bloom and Van Reenen (2007), adapted to the civil service following Rasul and Rogger (2017) in Nigeria and Rasul, Rogger, and Williams (forthcoming) in Ghana. See http://worldmanagementsurvey.org/ for further details on the World Management Survey instruments. BVR refers to Bloom and Van Reenen (2007).

Table A4: Services and supervisor bias at the organisation level

OLS Estimates

The dependent variable is the z-score of the service delivery indicator across all policy domains

Robust standard errors in all columns

	(1) Unconditional	(2) Sector Fixed effects	(3) Employee characteristics	(4) Woreda characteristics	(5) Management
Bias	-0.045	-0.047	-0.042	-0.061	-0.060
	[0.027]	[0.027]*	[0.027]	[0.026]**	[0.025]**
Average performance score	-0.16	-0.17	-0.16	-0.15	-0.14
	[0.032]***	[0.032]***	[0.034]***	[0.032]***	[0.035]***
Sector fixed effects	No	Yes	Yes	Yes	Yes
Employee characteristics	No	No	Yes	Yes	Yes
Jurisdiction characteristics	No	No	No	Yes	Yes
Management practices	No	No	No	No	Yes
Manager characteristics	No	No	No	No	Yes
Standard deviation in rainfall	No	No	No	No	No
Manager performance score	No	No	No	No	No
Adjusted R-squared	0.15	0.12	0.13	0.31	0.34
Observations	118	118	118	118	118

Notes: *** denotes significance at 1%, ** at 5%, and * at 10% level. All columns show OLS estimates. Robust standard errors are shown in all columns. The unit of observation is at the organisation level. The dependent variable is the organisation-average across all z-scores of the service-delivery indicators, where the z-score is calculated within indicator and across jurisdictions. The analysis is restricted to organisations for which there is BSC evaluation data. Bias is the absolute difference between the z score of the performance score and the z score of the attitude score. Positive bias is the interaction between the absolute difference between the z score of the performance score and the z score of the attitude score and an indicator for whether the actual difference is positive (z_attitude-z_performance>=0); Negative bias is similar, with an interaction for whether the actual difference is negative (z_attitude-z_performance<0). Individual characteristics include the organisation-average values for: the education level of the employees; an indicator for whether the employee is female; the average effort exerted by employees; the tenure in the position; and the tenure in the civil service. Effort is measured by Hours worked during a typical week is the respondent's response to "What is the actual number of hours you work in the civil service in a typical week?". Management practices refer to the aggregate management practices index based on the World Management Survey (Bloom and Van Reenen, 2007; Bloom et al 2012). Management characteristics include organisation-level averages of the tenure in the organisation of the management, tenure in the civil service of management, the education level of the management, the number of principals, and the extent of multitasking experienced by employeed on average. Sector fixed effects are indicators for whether the respondent works in an organisation that is within the agriculture, education, health, revenue, or trade sector. Jurisdiction characteristics include the population of the jurisdiction; the percentage of rural inhabitants; the ethnic fractionalisation (in z score); the remoteness measured in terms of the average travel time to the nearest urban centre of at least 50,000 people (a census measure); the road density in the jurisdiction; and the organisation-average difference in service delivery between the most recent two periods of data (averaged across all indicators, for a measure of overall uncertainty). Ethnic fractionalisation measures the probability that two randomly selected individuals in the jurisdiction are from different ethnicities (Fearon, 2003). The standard deviation in rainfall across the year uses the full set of observations for 2015 from the World Health Organisation LEAP programme and calculates the standard deviation. For a subset of organisations, we also have the performance scores of the managers, as evaluated by the head of organisation. Figures are rounded to two significant figures.

Table A5: Services and positive and negative supervisor bias. Conditions in which bias is more or less costly
OLS Estimates
The dependent variable is the z-score of the service delivery indicator across all policy domains
Standard errors clustered at the organisation level in all columns

	(1) Baseline	(2) Ethnic fractionalisation	(3) Linguistic fractionalisation	(4) Number of principals (monitoring)	(5) Monitoring technology
Positive bias	-0.073	-0.12	-0.12	-0.20	-0.0042
	[0.031]**	[0.030]***	[0.028]***	[0.069]***	[0.039]
Negative bias	-0.0011	-0.029	-0.011	-0.18	0.058
	[0.043]	[0.044]	[0.041]	[0.083]**	[0.049]
Average performance score	-0.18	-0.19	-0.20	-0.17	-0.17
	[0.042]***	[0.044]***	[0.043]***	[0.042]***	[0.041]***
Ethnic fractionalisation	0.033	0.14	0.31	0.030	0.045
	[0.042]	[0.052]***	[0.18]*	[0.042]	[0.040]
Positive bias x Ethnic fractionalisation		-0.11			
		[0.024]***			
Negative bias x Ethnic fractionalisation		-0.056			
		[0.075]			
Linguistic fractionalisation			-0.17		
			[0.17]		
Positive bias x Linguistic fractionalisation			-0.095		
			[0.024]***		
Negative bias x Linguistic fractionalisation			-0.032		
			[0.065]		
Number of principals				-0.076	
				[0.050]	
Positive bias x Number of principals				0.057	
				[0.031]*	
Negative bias x Number of principals				0.066	
				[0.026]**	
Inefficiency of monitoring technology					0.00021
					[0.0011]
Positve bias x Inefficiency of monitoring technology					-0.0014
					[0.00067]**
Negative bias x Inefficiency of monitoring technology					-0.0014
					[0.00074]*
Indicator fixed effects	Yes	Yes	Yes	Yes	Yes
Employee characteristics	Yes	Yes	Yes	Yes	Yes
Jurisdiction characteristics	Yes	Yes	Yes	Yes	Yes
Management practices	Yes	Yes	Yes	Yes	Yes
Manager characteristics	Yes	Yes	Yes	Yes	Yes
Standard deviation in rainfall	No	No	No	No	No
Manager performance score	No	No	No	No	No
Adjusted R-squared	0.10	0.10	0.10	0.10	0.11
Observations	1650 [118]	1650 [118]	1650 [118]	1650 [118]	1635 [117]

Notes: *** denotes significance at 1%, ** at 5%, and * at 10% level. All columns show OLS estimates. Standard errors clustered at the organisation level in all columns. The unit of observation is at the organisation-indicator level. The dependent variable is the z score of the service-delivery indicator, where the z score is defined within the indicator and across jurisdictions. The analysis is restricted to organisations for which there is BSC evaluation data. Bias is the absolute difference between the z score of the performance score and the z score of the attitude score. Positive bias is the interaction between the absolute difference between the z score of the performance score and the z score of the attitude score and an indicator for whether the actual difference is positive (z_attitude-z_performance>=0); Negative bias is similar, with an interaction for whether the actual difference is negative (z_attitude-z_performance<0). Individual characteristics include the organisation-average values for: the education level of the employees; an indicator for whether the employee is female; the average effort exerted by employees; the tenure in the position; and the tenure in the civil service. Effort is measured by Hours worked during a typical week is the respondent's response to "What is the actual number of hours you work in the civil service in a typical week?". Management practices refer to the aggregate management practices index based on the World Management Survey (Bloom and Van Reenen, 2007; Bloom et al 2012). Management characteristics include organisation-level averages of the tenure in the organisation of the management, tenure in the civil service of management, the education level of the management, the number of principals, and the extent of multitasking experienced by employeed on average. Sector fixed effects are indicators for whether the respondent works in an organisation that is within the agriculture, education, health, revenue, or trade sector. Jurisdiction characteristics include the population of the jurisdiction; the percentage of rural inhabitants; the ethnic fractionalisation (in z score); the remoteness measured in terms of the average travel time to the nearest urban centre of at least 50,000 people (a census measure); the road density in the jurisdiction; and the organisation-average difference in service delivery between the most recent two periods of data (averaged across all indicators, for a measure of overall uncertainty). Ethnic fractionalisation measures the probability that two randomly selected individuals in the jurisdiction are from different ethnicities (Fearon, 2003). Number of principals is the organisation average response to "How many people would you say regularly give you tasks as part of your formal work duties?". Inefficiency of monitoring technology is the organisation average response to "If you needed to request statistical data from the MIS (Management Information System) Directorate, how long do you expect it would take to get (assuming the data exists)?". Figures are rounded to two significant figures.

Table A6: Percentage of approved budget spent and organisation-level bias

OLS Estimates

The dependent variable is the percentage of the total approved budget of the woreda that is actually spent

Standard errors clustered at the jurisdiction level

	(1) Bias and budget unconditional	(2) Sector Fixed effects	(3) Budget size	(4) Performance scores	(5) Average employee characteristics	(6) Management	(7) Woreda characteristics
Bias	-0.00071	-0.00071	-0.0010	-0.0032	-0.0021	-0.0013	-0.00048
	[0.0025]	[0.0027]	[0.0029]	[0.0029]	[0.0027]	[0.0029]	[0.0030]
Approved budget (ETB millions)			-0.00033	-0.00032	-0.00047	-0.00067	-0.0011
			[0.00072]	[0.00067]	[0.00059]	[0.00059]	[0.00075]
Average performance score					-0.0058	-0.0052	-0.0019
					[0.0049]	[0.0040]	[0.0037]
Organisation service delivery performance							-0.015
							[0.0082]*
Sector fixed effects	No	No	Yes	Yes	Yes	Yes	Yes
Employee characteristics	No	No	No	No	Yes	Yes	Yes
Management practices	No	No	No	No	No	Yes	Yes
Manager characteristics	No	No	No	No	No	Yes	Yes
Jurisdiction characteristics	No	No	No	No	No	No	Yes
Adjusted R-squared	-0.0075	-0.041	-0.031	-0.0083	0.065	0.11	0.29
Observations	124 [35]	124 [35]	124 [35]	124 [35]	124 [35]	124 [35]	124 [35]

Notes: *** denotes significance at 1%, ** at 5%, and * at 10% level. All columns show OLS estimates. Standard errors are clustered at the jurisdiction level. The unit of observation is at the organisation-year level for the budget data. The dependent variable is the percentage of the total approved budget of the jurisdicition that is actually spent within the fiscal year for the year 2011/12 -- the latest year available for the budget data. The analysis is restricted to organisations for which there is BSC evaluation data and financial expenditure data (from BOOST). Bias is the absolute difference between the z score of the performance score and the z score of the attitude score. Positive bias is the interaction between the absolute difference between the z score of the performance score and the z score of the attitude score and an indicator for whether the actual difference is positive (z_attitude-z_performance>=0); Negative bias is similar, with an interaction for whether the actual difference is negative (z_attitude-z_performance<0). Individual characteristics include the organisation-average values for: the education level of the employees; an indicator for whether the employee is female; the average effort exerted by employees; the tenure in the position; and the tenure in the civil service. Effort is measured by Hours worked during a typical week is the respondent's response to "What is the actual number of hours you work in the civil service in a typical week?". Management practices refer to the aggregate management practices index based on the World Management Survey (Bloom and Van Reenen, 2007; Bloom et al 2012). Management characteristics include organisation-level averages of the tenure in the organisation of the management, tenure in the civil service of management, the education level of the management, the number of principals, and the extent of multitasking experienced by employeed on average. Sector fixed effects are indicators for whether the respondent works in an organisation that is within the agriculture, education, health, revenue, or trade sector. Jurisdiction characteristics include the population of the jurisdiction; the percentage of rural inhabitants; the ethnic fractionalisation (in z score); the remoteness measured in terms of the average travel time to the nearest urban centre of at least 50,000 people (a census measure); the road density in the jurisdiction; and the organisation-average difference in service delivery between the most recent two periods of data (averaged across all indicators, for a measure of overall uncertainty). Ethnic fractionalisation measures the probability that two randomly selected individuals in the jurisdiction are from different ethnicities (Fearon, 2003). Organisation service delivery performance is the organisation-average of the z-scores of the service delivery indicators, where the z-score of the service delivery indicators are constructed within indicator and across jurisdictions. Figures are rounded to two significant figures.

Figure 34: Correlation Between Attitude and Performance Evaluation Scores

Fig 1A: Correlation between attitude and performance scores

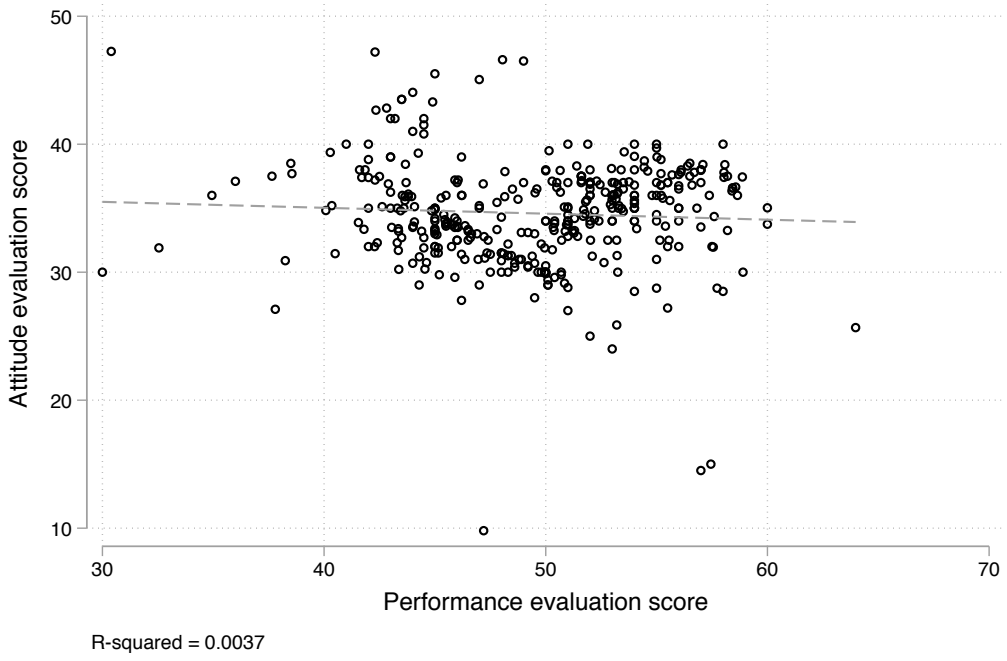
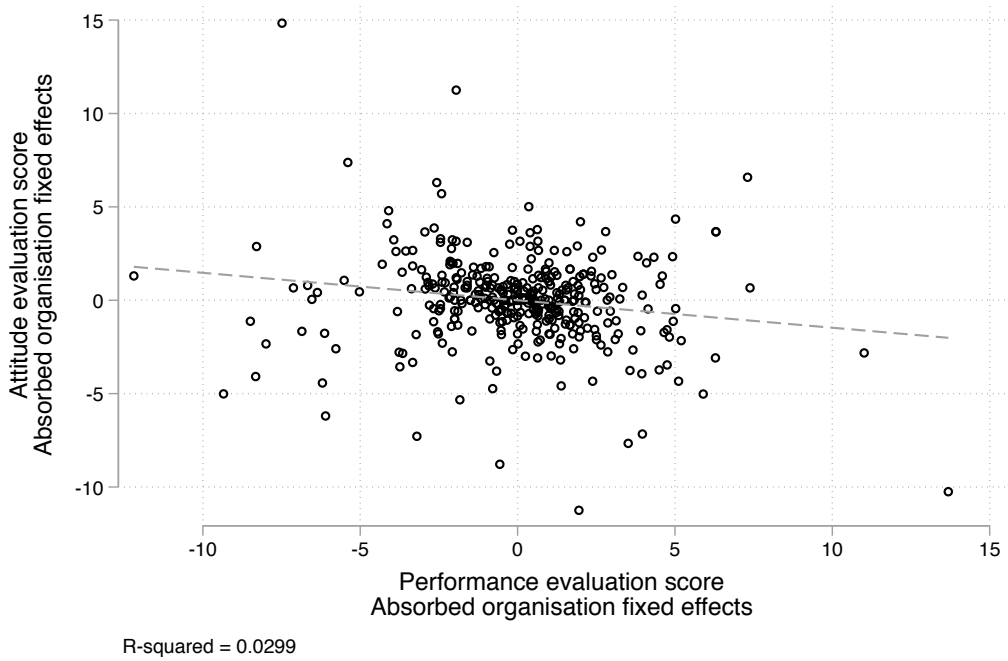
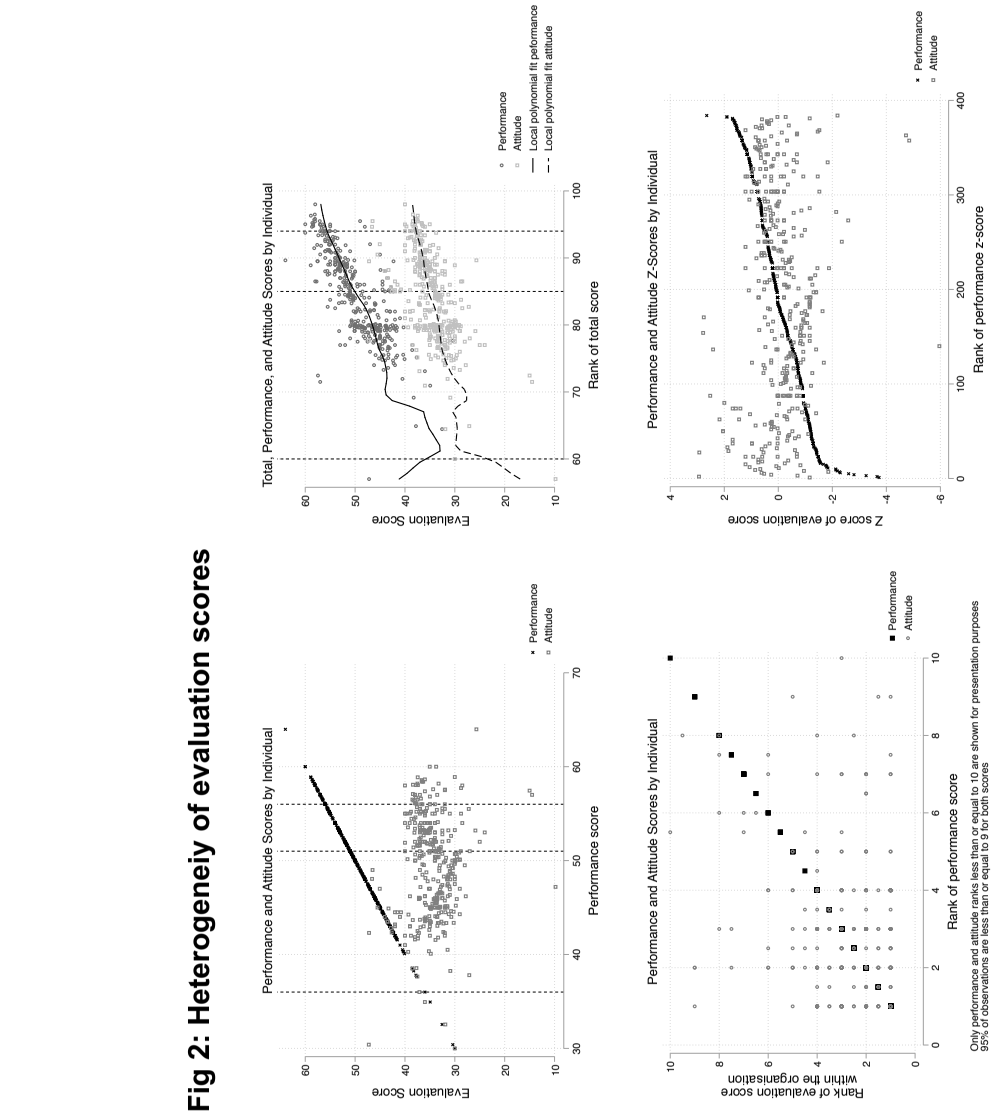


Fig 1B: Correlation between attitude and performance scores partialling out organisation fixed effects



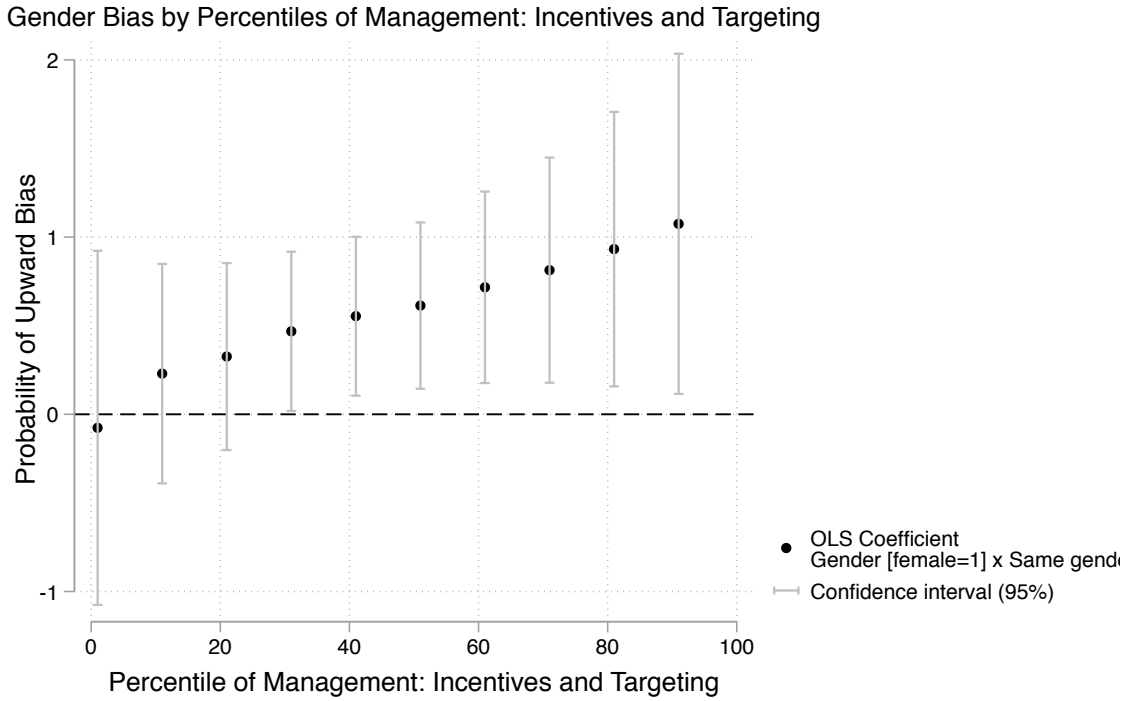
Notes: The top panel shows the unconditional correlation between the raw performance and raw attitude evaluation scores. The unit of observations is the civil servant. The R-Squared reported is form the unconditional regression of the attitude score on the performance score. The bottom panel shows the correlation between the raw performance and raw attitude evaluation scores, having accounted for organisation fixed effects (the residuals after running a regression of the score on the organisation fixed effects). The unit of observations is the civil servant. The R-Squared reported is from the unconditional regression of the residuals of the performance score and the attitude score having accounted for organisation fixed effects.

Figure 35: Individual-Level Variation in Attitude Scores and Performance Scores



Notes: The top-left figure shows the raw performance score and the raw attitude score for individuals ranked by their performance scorel the dashed lines represent cut offs for the performance score: "outstanding" is 56 and above; "very good" is 51 to 56; "satisfactory" is 36 to 51; "poor performance" is below 36. The top-right figure shows the raw performance score and raw attitude score for individuals ranked by their total evaluation score; the dashed lines represent cut offs for "outstanding" (94 or above), very good (between 85 and 94); and satisfactory (between 60 ad 85); below 60 is poor performace. The bottom-left figure shows the within-organisation rank of the performance and attitude score for individuals with the same ranked performance score; only the ranks of 10 or below are shown for presentation purposes and since this covers 95% of observations. The bottom-right figure shows the z-scores of the performance and attitude scores for individuals with the same ranked performance z-score. The unit of observation is the individual civil servant. The sample includes only those civil servants for which there is evaluation data.

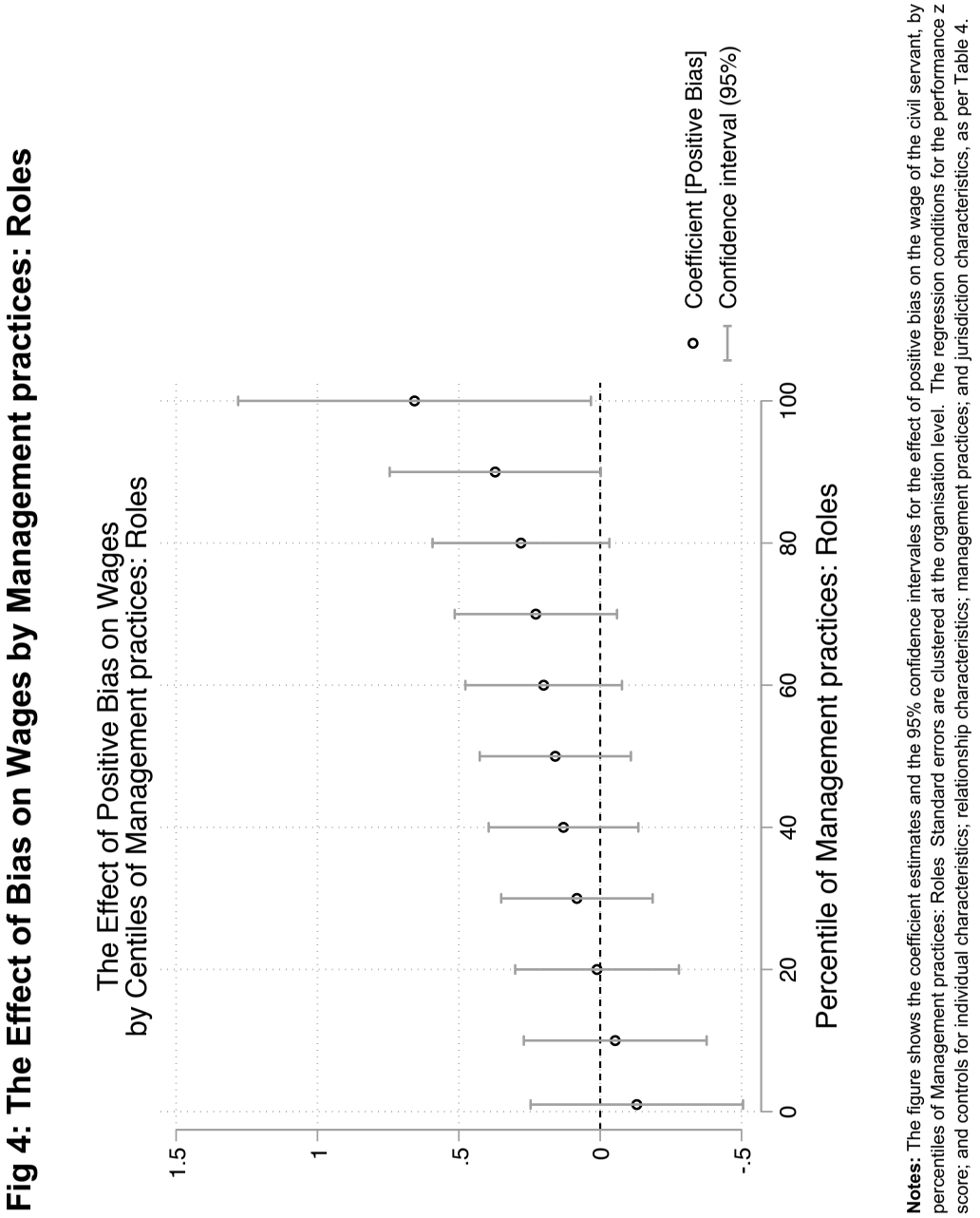
Fig 3: The Effect of Same Gender if Female on Upward Bias



Notes: The figure shows the coefficient estimates and the 95% confidence intervales for the effect of whether both the employee and manager are female on the probability of upward bias in the BSC evaluations of the civil servant, by percentiles of Management practices: Incentives and Targeting. Standard errors are clustered at the organisation level. The regression conditions for individual characteristics, sector fied effects, jurisdiction fixed effects, and management practices.

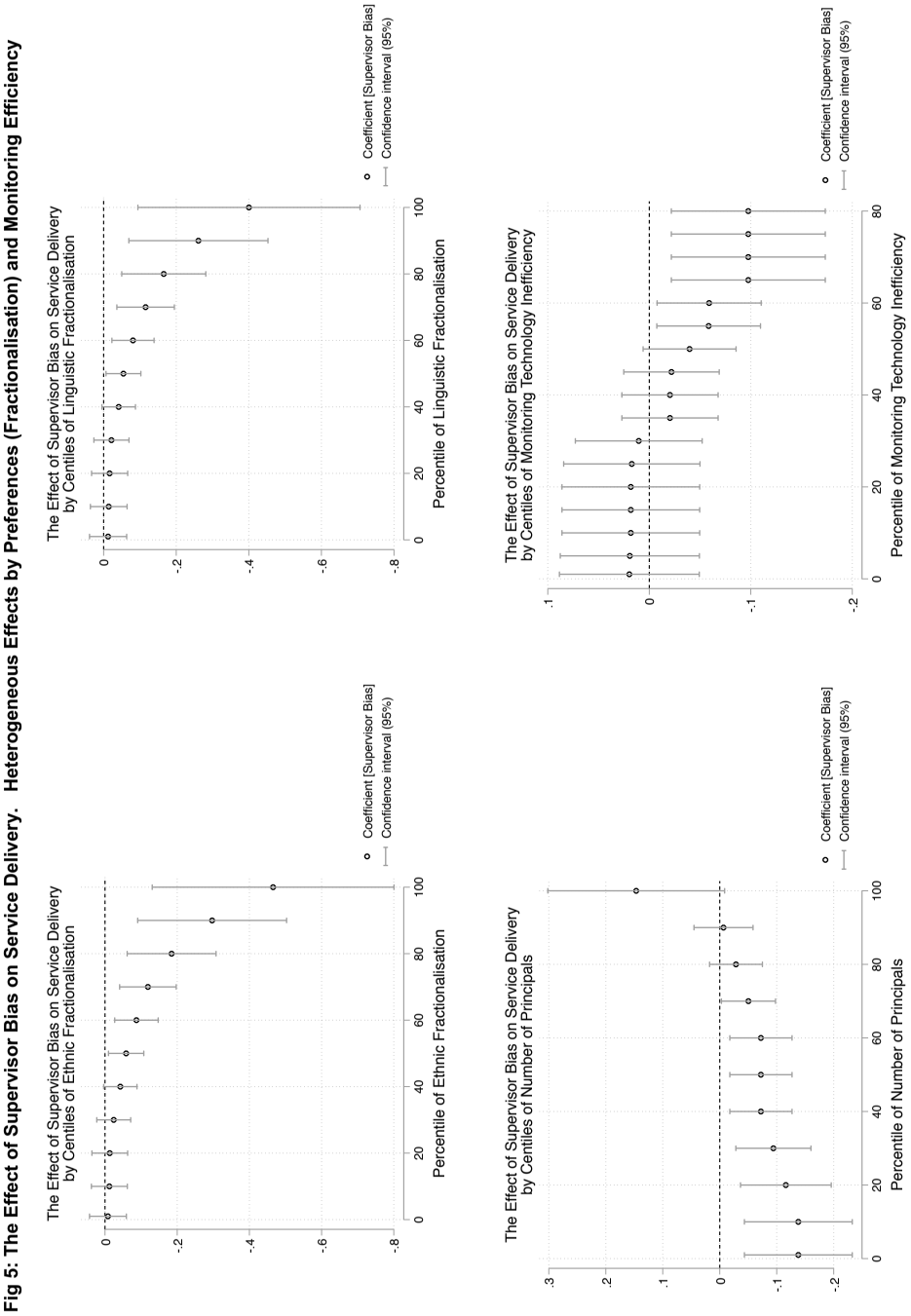
Figure 36: Same Gender and Supervisor Bias

Figure 37: Supervisor Bias and Wages by Management Practices Relating to Managerial Autonomy



Notes: The figure shows the coefficient estimates and the 95% confidence intervales for the effect of positive bias on the wage of the civil servant, by percentiles of Management practices: Roles Standard errors are clustered at the organisation level. The regression conditions for the performance z score; and controls for individual characteristics; relationship characteristics; management practices; and jurisdiction characteristics, as per Table 4.

Figure 38: Organisational Supervisor Bias and Service Delivery, by Local Environment Conditions



Notes: The top-left figure shows the regression coefficient and 95% confidence intervals for the effect of organisation level bias on the z-score of the service delivery indicators by centiles of ethnic fractionalisation. The top-right figure shows the regression coefficient and 95% confidence intervals for the effect of organisation level bias on the z-score of the service delivery indicators by centiles of linguistic fractionalisation. The bottom-left figure shows the regression coefficient and 95% confidence intervals for the effect of organisation level bias on the z-score of the service delivery indicators by centiles of the number of principals in the organisation. The bottom-right figure shows the regression coefficient and 95% confidence intervals for the effect of organisation level bias on the z-score of the service delivery indicators by centiles of monitoring technology. All regressions are conditional on sector fixed effects; individual characteristics of employees; management practices; manager characteristics; and jurisdiction characteristics. Standard errors are robust. The specification is equivalent to Table 6 column (5).

Fig A1A: Example Balanced Scorecard Performance Matrix

የግለሰብ ፈፃሚ 2008 ዓ/ም ስኮር ካርድ እቅድ አፈጻጸም ውጤት

ስትራተጂክ ግቦችን የሚያሳኩ መለኪያዎች(ግቦች)	ግቡን ለማስፈፀም የሚከናወኑ ዋና ዋና ውጤት ተኮር ተግባራት	ለውጤት ተኮር ተግባር ከመቶ የተሰጠው ክብደት በ %	የመቶኛ ድርሻ በ%	መለኪያ	መለኪያው ለፈፃሚው የሚደርሰው የመቶኛ ድርሻ በ%	2008 ኢላማ	2ኛ ሩብ ዓመትኢላማ	2ኛ ሩብ ዓመት አፈጻጸም ውጤት	ድምር ውጤት
ግብ1፡- በከተማ አስተዳሩ አሰራር ተፈጥሮላቸው ሞዴል የሆኑ የ1ለ5 ቡድኖች ከ46% ወደ 70% ማድረስ	**ተግባር 1** ቡድኑ በሚደግፋቸው ተቋማት የ1ለ5 ቡድን አደረጃጀቶች እንዲጠናከሩ ክትትልና ድጋፍ ማድረግ	**35**	**9**	መጠን/ድግግሞሽ/	35%	56	18	3.15	7.87
				ጊዜ(በሰዓት)	35%	56x0.42=23.5	18x0.42=7	3.15	
				ጥራት	30%	100	100	1.57	
	ተግባር 2 የየተቋማትን የ1ለ5 አደረጃጀቶች ያሉበትን ሁኔታ መረጃ አደራጅቶ መያዝ	**35**	9	መጠን/ድግግሞሽ/	35%	4	4	3.15	7.87
				ጊዜ(በሰዓት)	35%	4x8=32	4x8 =32	3.15	
				ጥራት	30%	100%	100%	1.57	7.87
	ተግባር 3. በመስፈርቱ መሰረት በተቋማት ሞዴል የ1ለ5 አደረጃጀቶች እንዲለዩ መደገፍና መረጃውን አደራጅቶ መያዝ	30	8	መጠን/ድግግሞሽ/	35%	56	18	3.15	
				ጊዜ(በሰዓት)	35%	56x0.42=23.5	18x0.42=7	3.15	7.87
				ጥራት	30%	100	100	1.57	
ግብ2፡- በከተማ አስተዳሩ አሰራር ተፈጥሮላቸው	**ተግባር1** ቡድኑ በሚደግፋቸው ተቋማት የለውጥ	35	9	መጠን/ድግግሞሽ/	35%	56	18	3.15	7.87
				ጊዜ(በሰዓት)	35%	56x0.42=23.5	18x0.42=7	3.15	

Figure 39: Example of Civil Servant Performance Evaluation Matrix (Page 1)

Fig A1B: Example Balanced Scorecard Performance Matrix (Continued)

ስትራተጂክ ግቦችን የሚያሳኩ መለኪያዎች(ግቦች)	ግቡን ለማስፈፀም የሚከናወኑ ዋና ዋና ውጤት ተኮር ተግባራት	ለውጤት ተኮር ተግባር ከመቶ የተሰጠው ክብደት በ %	የመቶኛ ድርሻ በ%	መለኪያ	መለኪያው ለፈፃሚው የሚደርሰው የመቶኛ ድርሻ በ%	2008 ኢላማ	2ኛ ሩብ ዓመትዒላማ	2ኛ ሩብ ዓመት አፈጻጸም ውጤት	ድምር ውጤት
ግብ 2፡-ሰራዊት የነኑቡ የማዕከል ሴክተር ተቋማት በመቶኛ	**ተግባር 1** ተቋማት በልማት ሰራዊት ግንባታ መመሪያ መስፈርት መሰረት እንዲሄዱ ለማስቻል በአደረጃጀቶች አፈጻጸምን እየፈተሹ የማጠናከር ስራ እንዲሰራ ክትትልና ድጋፍ ማድረግ፣	50	12	መጠን/ድግግሞሽ/	30%	56	18	4.2	10.5
				ጊዜ(በሰዓት)	30%	56x0.42=23.5	18x0.42=7	4.2	
				ጥራት	30%	100	100	2.1	
	ተግባር 2 ተቋማት ለፈጻሚዎች ተልዕኮ ሰጥቶ በማሰማራትና ውጤታቸውን በየጊዜው እየገመገሙ በመደገፍ ሰራዊት እንዲፈጠር ሚያስችል አቅጣጫ እንዲከተሉ ድጋፍ ክትትል ማድረግ፣	50	12	መጠን/ድግግሞሽ	30%	56	18	4.2	10.5
				ጊዜ(በሰዓት)	30%	56x0.42=23.5	18x0.42=7	4.2	
				ጥራት	30%	100	100	2.1	
ግብ 2፡- የተፈጠሩ የእርስ በእርስ የመማማሪያ መድረኮች ብዛት	**ተግባር 1**፡- የእርስ በእርስ የመማሪያ መድረኮች በየሳምንቱ ማካሄድ	35	8	መጠን/ድግግሞሽ/	30%	36	12	2.81	7.02
				ጊዜ(በሰዓት)	30%	36x1=36	12	2.81	
				ጥራት	30%	100	100	1.4	
	ከ 417								353.71
	ከ70% ስኮር ካርድ								59.37
	ከ20% ዙሪያ መለስ								18
	ከ10 % ራስን ማብቃት								9
	ከ100%								86.37

ነጥቡን የሞላው የኬዝ ቲሙ አስተባበሪ ስም ---------- ፈጻሚ ባለሙያ ስም ----------

ፊርማ ----------

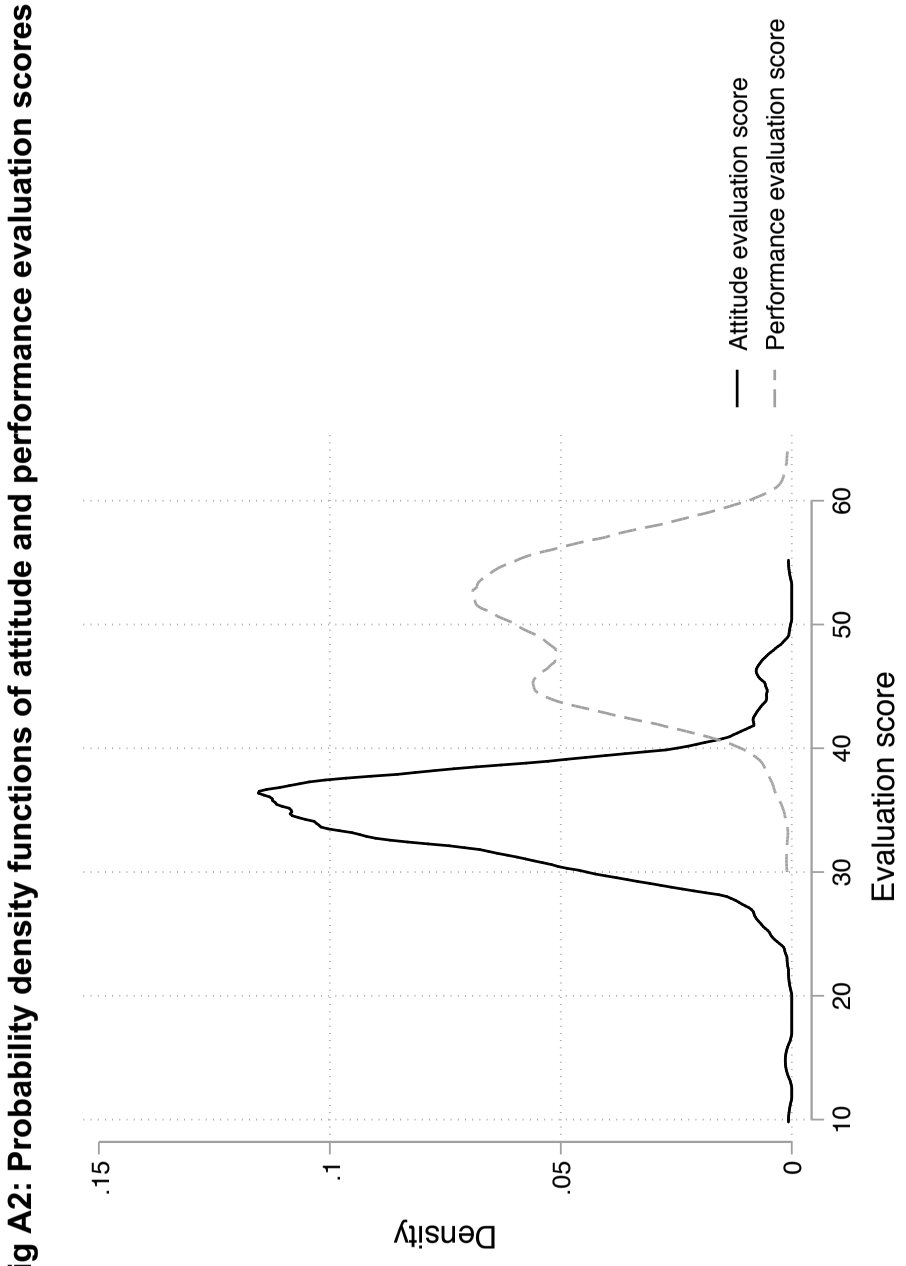
ቀን ----------

ፊርማ ----------

ቀን ----------

Figure 40: Example of Civil Servant Performance Evaluation Matrix (Page 2)

Figure 41: Distributions of Evaluation Scores Across Individuals



Notes: The figure shows the probability density functions of the attitude and performance score across all civil servants in the Ethiopian Civil Servants Survey for whom there is Balanced Scorecard data.

References

Abagisa, J. [2014], Challenges and prospects of Balanced scorecard Implementation in the Development Bank of Ethiopia, PhD thesis, Addis Ababa University.

Abebe, G., Caria, S., Fafchamps, M., Falco, P., Franklin, S. and Quinn, S. [2016], Curse of anonymity or tyranny of distance? the impacts of job-search support in urban ethiopia, Working Paper 22409, National Bureau of Economic Research.
URL: *http://www.nber.org/papers/w22409*

Acemoglu, D., Aghion, P., Lelarge, C., Van Reenen, J. and Zilibotti, F. [2007], 'Technology, information, and the decentralization of the firm', *Quarterly Journal of Economics* **122**(4), 1759–1799.

Acemoglu, D., Garcia-Jimeno, C. and Robinson, J. A. [2015], 'State capacity and economic development: A network approach', *American Economic Review* **105**(8), 2364–2409.

Acemoglu, D. and Robinson, J. [2012*a*], 'Why nations fail', *New York: Crown Business* .

Acemoğlu, D. and Robinson, J. A. [2012*b*], 'Why nations fail: The origins of power, prosperity, and poverty', *New York: Crown Business* .

Acemoglu, D. and Zilibotti, F. [2001], 'Productivity differences', *The Quarterly Journal of Economics* **116**(2), 563–606.

Addis Fortune [2014], 'Government announces scale of civil servant salary increment'.

Agarwal, S. and Ben-David, I. [2018], 'Loan prospecting and the loss of soft information', *Journal of Financial Economics* .

Agarwal, S. and Hauswald, R. [2010*a*], 'Distance and private information in lending', *The Review of Financial Studies* **23**(7), 2757–2788.

Agarwal, S. and Hauswald, R. B. [2010*b*], 'Authority and information'.

Aghion, P., Bloom, N., Lucking, B., Sadun, R. and Van Reenen, J. [2017], Turbulence, firm decentralization and growth in bad times, Technical report, National Bureau of Economic Research.

Aghion, P. and Tiróle, J. [1997], 'Formal and real authority in organizations', *Journal of Political Economy* **105**(1), 1–29.

Aker, J. C., Boumnijel, R., McClelland, A. and Tierney, N. [2016], 'Payment mechanisms and antipoverty programs: Evidence from a mobile money cash transfer experiment in niger', *Economic Development and Cultural Change* **65**(1), 1–37.

Albrecht, J., Robayo-Abril, M. and Vroman, S. [2015], 'Public-sector employment in an equilibrium search and matching model', *The Economic Journal* **doi:10.1111/ecoj.12533.**

Andrade, P. and Le Bihan, H. [2013], 'Inattentive professional forecasters', *Journal of Monetary Economics* **60**(8), 967–982.

Argenziano, R., Severinov, S. and Squintani, F. [2016], 'Strategic information acquisition and transmission', *American Economic Journal: Microeconomics* **8**(3), 119–55.

Ashcroft, K. and Rayner, P. [2011], *Higher Education in Development: Lessons from Sub Saharan Africa*, IAP.

Ashenfelter, O. [1978], 'Estimating the effect of training programs on earnings', *The Review of Economics and Statistics* pp. 47–57.

Ashraf, N., Bandiera, O. and Jack, B. K. [2014], 'No margin, no mission? a field experiment on incentives for public service delivery', *Journal of Public Economics* **120**, 1–17.

Ashraf, N., Bandiera, O. and Lee, S. S. [2014], 'Awards unbundled: Evidence from a natural field experiment', *Journal of Economic Behavior & Organization* **100**, 44–63.

Ashraf, N., Bandiera, O. and Lee, S. S. [2016], 'Do-gooders and go-getters'.

Atkin, D. [2016], 'Endogenous skill acquisition and export manufacturing in mexico', *The American Economic Review* **106**(8), 2046–2085.

Baker, G., Gibbons, R. and Murphy, K. J. [1994], 'Subjective performance measures in optimal incentive contracts', *The Quarterly Journal of Economics* **109**(4), 1125–1156.

Bandiera, O., Best, M., Khan, A. and Prat, A. [2018], 'Incentives and the allocation of authority in organizations: A field experiment with bureaucrats', *Mimeo: London School of Economics* .

Banerjee, A., Hanna, R., Kyle, J. C., Olken, B. A. and Sumarto, S. [2015], The power of transparency: information, identification cards and food subsidy programs in indonesia, Technical report, National Bureau of Economic Research.

Banerjee, A., Hanna, R., Kyle, J., Olken, B. A. and Sumarto, S. [2018], 'Tangible information and citizen empowerment: Identification cards and food subsidy programs in indonesia', *Journal of Political Economy* **126**(2), 451–491.

Banerjee, A. V. and Duflo, E. [2005], 'Growth theory through the lens of development economics', *Handbook of economic growth* **1**, 473–552.

Bardhan, P. [2002], 'Decentralization of governance and development', *Journal of Economic perspectives* **16**(4), 185–205.

Bardhan, P. and Mookherjee, D. [2006], 'Decentralisation and accountability in infrastructure delivery in developing countries', *The Economic Journal* **116**(508), 101–127.

Barnwal, P. [2014], 'Curbing leakage in public programs with biometric identification systems: Evidence from india's fuel subsidies', *Manuscript. I* .

Barro, R. J. [2001], 'Human capital and growth', *American economic review* **91**(2), 12–17.

Basinga, P., Gertler, P. J., Binagwaho, A., Soucat, A. L., Sturdy, J. and Vermeersch, C. M. [2011], 'Effect on maternal and child health services in rwanda of payment to primary health-care providers for performance: an impact evaluation', *The Lancet* **377**(9775), 1421–1428.

Bender, S., Bloom, N., Card, D., Van Reenen, J. and Wolter, S. [2018], 'Management practices, workforce selection, and productivity', *Journal of Labor Economics* **36**(S1), S371–S409.

Bentley MacLeod, W. [2003], 'Optimal contracting with subjective evaluation', *The American Economic Review* **93**(1), 216–240.

Bernheim, B. D. and Whinston, M. D. [1986], 'Common agency', *Econometrica: Journal of the Econometric Society* pp. 923–942.

Bertrand, M., Burgess, R., Chawla, A. and Xu, G. [2016], 'The costs of bureaucratic rigidity: Evidence from the indian administrative service', *Unpublished working paper. University of Chicago* .

Bertrand, M., Burgess, R., Chawla, A. and Xu, G. [2017], The glittering prizes: Career incentives and bureaucrat performance, Technical report, mimeo.

Bertrand, M., Duflo, E. and Mullainathan, S. [2004], 'How much should we trust differences-in-differences estimates?', *The Quarterly journal of economics* **119**(1), 249–275.

Besley, T. and Ghatak, M. [2005], 'Competition and incentives with motivated agents', *American Economic Review* **95**(3), 616–636.

Besley, T. and Persson, T. [2010], 'State capacity, conflict, and development', *Econometrica* **78**(1), 1–34.

Besley, T. and Persson, T. [2011], *Pillars of Prosperity: The Political Economics of Development Clusters*, Princeton University Press.

Bloom, D. E., Canning, D. and Chan, K. [2006], *Higher Education and Economic Development in Africa*, Vol. 102, World Bank Washington, DC.

Bloom, N., Propper, C., Seiler, S. and Van Reenen, J. [2015], 'The impact of competition on management quality: Evidence from public hospitals', *The Review of Economic Studies* **82**(2), 457–489.

Bloom, N., Sadun, R. and Van Reenen, J. [2012], 'The organization of firms across countries', *The Quarterly Journal of Economics* **127**(4), 1663–1705.

Bloom, N. and Van Reenen, J. [2007*a*], 'Measuring and explaining management practices across firms and countries', *The Quarterly Journal of Economics* **122**(4), 1351–1408.

Bloom, N. and Van Reenen, J. [2007*b*], 'Measuring and explaining management practices across firms and countries*', *The Quarterly Journal of Economics* **122**(4), 1351–1408.
URL: *+ http://dx.doi.org/10.1162/qjec.2007.122.4.1351*

Bloom, N. and Van Reenen, J. [2010], 'New approaches to surveying organizations', *The American Economic Review* **100**(2), 105–109.

Bloom, N., Van Reenen, J. et al. [2010], 'Why do management practices differ across firms and countries?', *Journal of Economic Perspectives* **24**(1), 203–224.

Boccanfuso, D., Larouche, A. and Trandafir, M. [2015], 'Quality of higher education and the labor market in developing countries: Evidence from an education reform in senegal', *World Development* **74**, 412 – 424.
URL: *http://www.sciencedirect.com/science/article/pii/S0305750X15001163*

Boräng, F., Cornell, A., Grimes, M. and Schuster, C. [2018], 'Cooking the books: Bureaucratic politicization and policy knowledge', *Governance* **31**(1), 7–26.

Bradley, J., Postel-Vinay, F. and Turon, H. [2017], 'Public sector wage policy and labor market equilibrium: a structural model', *Journal of the European Economic Association* p. jvw026.

Brown, M., Schaller, M., Westerfeld, S. and Heusler, M. [2012], Information or insurance? on the role of discretion in relationship lending, Technical report, Working Paper, University of St. Gallen.

Brynjolfsson, E. and McElheran, K. [2016], 'The rapid adoption of data-driven decision-making', *American Economic Review* **106**(5), 133–39.

Burdett, K. [2012], 'Towards a theory of the labor market with a public sector', *Labour Economics* **19**(1), 68–75.

Callen, M., Gulzar, S., Hasanain, A., Khan, M. Y. and Rezaee, A. [2018], 'Data and policy decisions: Experimental evidence from pakistan'.

Callen, M., Gulzar, S., Hasanain, A., Khan, Y. and Rezaee, A. [2015], Personalities and public sector performance: Evidence from a health experiment in pakistan, Technical report, National Bureau of Economic Research.

Callen, M. and Long, J. D. [2014], 'Institutional corruption and election fraud: Evidence from a field experiment in afghanistan', *The American Economic Review* **105**(1), 354–381.

Cameron, A. C. and Miller, D. L. [2015], 'A practitioner's guide to cluster-robust inference', *Journal of Human Resources* **50**(2), 317–372.

Card, D. [2001], 'Estimating the return to schooling: Progress on some persistent econometric problems', *Econometrica* **69**(5), 1127–1160.

Card, D. and Lemieux, T. [2001], 'Can falling supply explain the rising return to college for younger men? a cohort-based analysis', *The Quarterly Journal of Economics* **116**(2), 705–746.

Casey, K. [2015], 'Crossing party lines: The effects of information on redistributive politics', *American Economic Review* **105**(8), 2410–48.
URL: *http://www.aeaweb.org/articles.php?doi=10.1257/aer.20130397*

Chamberlain, A. [2015], Why is hiring taking longer? new insights from glassdoor data, Technical report, Glassdoor.
URL: *https://www.glassdoor.com/research/app/uploads/sites/2/2015/06/GD_Report_3.pdf*

Chamberlain, A. [2017], Glassdoor economic research blog, Technical report, Glassdoor.
URL: *https://www.glassdoor.com/research/time-to-hire-in-25-countries/*

Chamberlain, G. [1984], Panel data, *in* G. Z. and D. I. M., eds, 'Handbook of Econometrics', 1 edn, Vol. 2, Elsevier, chapter 22, pp. 1247–1318.
URL: *http://EconPapers.repec.org/RePEc:eee:ecochp:2-22*

Cowley, E. and Smith, S. [2014], 'Motivation and mission in the public sector: evidence from the world values survey', *Theory and decision* **76**(2), 241–263.

Crawford, G. and Hartmann, C. [2008], *Decentralisation in Africa: a pathway out of poverty and conflict?*, Amsterdam University Press.

Crawford, V. P. and Sobel, J. [1982], 'Strategic information transmission', *Econometrica* **50**(6), 1431–1451.

Cremer, J., Estache, A. and Seabright, P. [1994], *The decentralization of public services: lessons from the theory of the firm*, Vol. 1345, World Bank Publications.

D'agostino, R. B., Belanger, A. and D'Agostino Jr, R. B. [1990], 'A suggestion for using powerful and informative tests of normality', *The American Statistician* **44**(4), 316–321.

Dal Bó, E., Finan, F. and Rossi, M. A. [2013], 'Strengthening state capabilities: The role of financial incentives in the call to public service*', *Quarterly Journal of Economics* **128**(3), 1169–1218.

De Meulemeester, J.-L. and Rochat, D. [1995], 'A causality analysis of the link between higher education and economic development', *Economics of Education Review* **14**(4), 351–361.

Deserranno, E. [2017], 'Financial incentives as signals: experimental evidence from the recruitment of village promoters in uganda'.

Dessein, W. [2002], 'Authority and communication in organizations', *The Review of Economic Studies* **69**(4), 811–838.

Dhaliwal, I. and Hanna, R. [2017], 'The devil is in the details: The successes and limitations of bureaucratic reform in india', *Journal of Development Economics* **124**, 1–21.

Di Pei, H. [2015], 'Communication with endogenous information acquisition', *Journal of Economic Theory* **160**, 132–149.

Dixit, A. [1997], 'Power of incentives in private versus public organizations', *The American Economic Review* **87**(2), 378–382.

Dixit, A. [2002*a*], 'Incentives and organizations in the public sector: An interpretative review', *Journal of Human Resources* **37**(4), 696–727.

Dixit, A. [2002*b*], 'Incentives and organizations in the public sector: An interpretative review', *The Journal of Human Resources* **37**(4), pp. 696–727.
URL: *http://www.jstor.org/stable/3069614*

Dorosh, P. and Schmidt, E. [2010], The rural-urban transformation in ethiopia, Technical report, Citeseer.

Duflo, E. [2001], 'Schooling and labor market consequences of school construction in indonesia: Evidence from an unusual policy experiment', *American economic review* **91**(4), 795–813.

Duflo, E. [2004], 'The medium run effects of educational expansion: Evidence from a large school construction program in indonesia', *Journal of Development Economics* **74**(1), 163–197.

Duflo, E. [2017], 'Richard t. ely lecture: The economist as plumber', *American Economic Review* **107**(5), 1–26.

Duggan, M. and Scott Morton, F. M. [2006], 'The distortionary effects of government procurement: evidence from medicaid prescription drug purchasing', *The Quarterly Journal of Economics* **121**(1), 1–30.

Ehrbeck, T. and Waldmann, R. [1996], 'Why are professional forecasters biased? agency versus behavioral explanations', *The Quarterly Journal of Economics* **111**(1), 21–40.

Engel, J. [2010], 'Ethiopia's progress on education: A rapid and equitable expansion of access'.

Ethiopian Civil Service University [2012], 'A training on educational management information system'.
URL: *http://www.ecsu.edu.et/article/training-educational-management-information-system*

Fan, C. S. and Stark, O. [2007], 'International migration and "educated unemployment"', *Journal of Development Economics* **83**(1), 76–87.

Federal Democratic Republic of Ethiopia [2003], Rural development policy and strategies, Technical report, Federal Democratic Republic of Ethiopia.

Federal Democratic Republic of Ethiopia [2010], Growth and transformation plan (2010/11-2014/15), Technical report, Federal Democratic Republic of Ethiopia.

Federal Democratic Republic of Ethiopia [2012], 'Ministries and agencies', Online.
URL: *http://www.ethiopia.gov.et/en_GB/ministries-and-agencies*

Federal Democratic Republic of Ethiopia [2016], Growth and transformation plan ii (2015/16-2019/20), Technical report, Federal Democratic Republic of Ethiopia.

Federal Ministry of Education [2015], Education sector development programme v (esdp v), Technical report, Federal Democratic Republic of Ethiopia.

Ferraz, C. and Finan, F. [2008], 'Exposing corrupt politicians: The effects of brazil's publicly released audits on electoral outcomes', *The Quarterly Journal of Economics* **123**(2), 703–745.
URL: *http://qje.oxfordjournals.org/content/123/2/703.abstract*

Filmer, D. [2007], 'If you build it, will they come? school availability and school enrolment in 21 poor countries', *The Journal of Development Studies* **43**(5), 901–928.

Finan, F., Olken, B. A. and Pande, R. [2017], The personnel economics of the developing state, *in* 'Handbook of Economic Field Experiments', Vol. 2, Elsevier, pp. 467–514.

Fisher, J. R., Montambault, J., Burford, K. P., Gopalakrishna, T., Masuda, Y. J., Reddy, S. M., Torphy, K. and Salcedo, A. I. [2018], 'Knowledge diffusion within a large conservation organization and beyond', *PloS one* **13**(3), e0193716.

Frederiksen, A., Kahn, L. B. and Lange, F. [2017], 'Supervisors and performance management systems', *National Bureau of Economic Research* .

Fujiwara, T. [2015], 'Voting technology, political responsiveness, and infant health: Evidence from brazil', *Econometrica* **83**(2), 423–464.

Gabaix, X., Laibson, D., Moloche, G. and Weinberg, S. [2006], 'Costly information acquisition: Experimental analysis of a boundedly rational model', *American Economic Review* **96**(4), 1043–1068.

Gant, J., Ichniowski, C. and Shaw, K. [2002], 'Social capital and organizational change in high-involvement and traditional work organizations', *Journal of Economics & Management Strategy* **11**(2), 289–328.

Garicano, L. [2000], 'Hierarchies and the organization of knowledge in production', *Journal of Political Economy* **108**(5), 874–904.

Gebre-Egziabher, T. [2007], *Decentralization in Ethiopia*, African Books Collective.

Georgellis, Y., Iossa, E. and Tabvuma, V. [2010], 'Crowding out intrinsic motivation in the public sector', *Journal of Public Administration Research and Theory* **21**(3), 473–493.

Girsberger, E. and Meango, R. [2017], 'The puzzle of educated unemployment in west africa', *Working Paper* .

Glewwe, P. [2002], 'Schools and skills in developing countries: Education policies and socioeocnomic outcomes', *Journal of Economic Literature* **40**(2), 436.

Glewwe, P., Ilias, N. and Kremer, M. [2010], 'Teacher incentives', *American Economic Journal: Applied Economics* **2**(3), 205–27.

Glewwe, P. and Kremer, M. [2006], 'Schools, teachers, and education outcomes in developing countries', *Handbook of the Economics of Education* **2**, 945–1017.

Goldfinch, S., DeRouen Jr, K. and Pospieszna, P. [2013], 'Flying blind? evidence for good governance public management reform agendas, implementation and outcomes in low income countries', *Public Administration and Development* **33**(1), 50–61.

Gomes, P. [2015], 'Optimal public sector wages', *The Economic Journal* **125**(587), 1425–1451.

Gomes, P. [2017], 'Heterogeneity and the public sector wage policy', *International Economic Review* .

Gulzar, S. and Pasquale, B. J. [2017], 'Politicians, bureaucrats, and development: Evidence from india', *American Political Science Review* **111**(1), 162–183.

Hanna, R., Mullainathan, S. and Schwartzstein, J. [2014], 'Learning through noticing: Theory and evidence from a field experiment', *The Quarterly Journal of Economics* **129**(3), 1311–1353.

Harris, R. D. and Tzavalis, E. [1999], 'Inference for unit roots in dynamic panels where the time dimension is fixed', *Journal of econometrics* **91**(2), 201–226.

Hart, O. and Holmstrom, B. [1987], The theory of contracts/advance in economic theory, *in* 'Tirth World Congress. Ed. by T. Bewlew. Cambridge University Press'.

Hastings, J. S. and Weinstein, J. M. [2008], 'Information, school choice, and academic achievement: Evidence from two experiments', *Quarterly Journal of Economics* **123**(4), 1373–1414.

Hattori, M., Shintani, K. and Uchida, H. [2015], 'The repository of soft information within bank organizations', *Journal of Money, Credit and Banking* **47**(4), 737–770.

Hörner, J., Ngai, L. R. and Olivetti, C. [2007], 'Public enterprises and labor market performance', *International Economic Review* **48**(2), 363–384.

Hsieh, C.-T. and Klenow, P. J. [2009], 'Misallocation and manufacturing tfp in china and india*', *The Quarterly Journal of Economics* **124**(4), 1403–1448.
URL: *http://dx.doi.org/10.1162/qjec.2009.124.4.1403*

Im, K. S., Pesaran, M. H. and Shin, Y. [2003], 'Testing for unit roots in heterogeneous panels', *Journal of Econometrics* **115**, 53–74.

Iyer, L. and Mani, A. [2012], 'Traveling agents: political change and bureaucratic turnover in india', *Review of Economics and Statistics* **94**(3), 723–739.

James, Q. W. [1989], 'Bureaucracy: What government agencies do and why they do it', *New York: Basic. Google Scholar* .

Jessoe, K. and Rapson, D. [2014], 'Knowledge is (less) power: Experimental evidence from residential energy use', *American Economic Review* **104**(4), 1417–38.

Khan, A. Q., Khwaja, A. I. and Olken, B. A. [2015], 'Tax farming redux: Experimental evidence on performance pay for tax collectors', *The Quarterly Journal of Economics* **131**(1), 219–271.

Khan, Q. M., Faguet, J.-P., Gaukler, C. and Mekasha, W. [2014], *Improving Basic Services for the Bottom Forty Percent: Lessons from Ethiopia*, The World Bank.
URL: *http://elibrary.worldbank.org/doi/abs/10.1596/978-1-4648-0331-4*

Laffont, J.-J. and Martimort, D. [1998], 'Collusion and delegation', *The Rand Journal of Economics* pp. 280–305.

Landier, A., Nair, V. B. and Wulf, J. [2007], 'Trade-offs in staying close: Corporate decision making and geographic dispersion', *The Review of Financial Studies* **22**(3), 1119–1148.

Levin, J. [2003], 'Relational incentive contracts', *The American Economic Review* **93**(3), 835–857.

Lewis-Faupel, S., Neggers, Y., Olken, B. A. and Pande, R. [2016], 'Can electronic procurement improve infrastructure provision? evidence from public works in india and indonesia', *American Economic Journal: Economic Policy* **8**(3), 258–83.

Li, S., Whalley, J. and Xing, C. [2014], 'China's higher education expansion and unemployment of college graduates', *China Economic Review* **30**, 567–582.

Liberti, J. M. [2017], 'Initiative, incentives, and soft information', *Management Science* .

MacLeod, W. B. and Malcomson, J. M. [1989], 'Implicit contracts, incentive compatibility, and involuntary unemployment', *Econometrica: Journal of the Econometric Society* pp. 447–480.

Mankiw, N. G., Romer, D. and Weil, D. N. [1992], 'A contribution to the empirics of economic growth*', *The Quarterly Journal of Economics* **107**(2), 407–437.
URL: *http://dx.doi.org/10.2307/2118477*

Markos, S. [2013], 'Civil service reform in ethiopia: Issues, lessons, and future directions', *International Journal of Public Administration* **36**(4), 235–247.
URL: *https://doi.org/10.1080/01900692.2012.713305*

Martimort, D. [1996], 'The multiprincipal nature of government', *European Economic Review* **40**(3), 673–685.

Mas-Collell, A., Whinston, M. and Green, J. R. [1995], 'Microeconomic theory'.

Masaki, T., Custer, S., Eskenazi, A., Stern, A. and Latourell, R. [2017], 'Decoding data use: How do leaders use data and use it to accelerate development', *Williamsburg, VA: AidData at the College of William & Mary. ii* .

Melumad, N., Mookherjee, D. and Reichelstein, S. [1992], 'A theory of responsibility centers', *Journal of Accounting and Economics* **15**(4), 445–484.

Meng, X. and Gregory, R. G. [2002], 'The impact of interrupted education on subsequent educational attainment: A cost of the chinese cultural revolution', *Economic Development and Cultural Change* **50**(4), 935–959.
URL: *https://doi.org/10.1086/342761*

Meyers, M. K. and Vorsanger, S. [2007], 'Street-level bureaucrats and the implementation of public policy', *The Handbook of Public Administration* pp. 153–163.

Miller, A. R. and Tucker, C. [2011], 'Can health care information technology save babies', *Journal of Political Economy* **119**(2), 289–324.

Ministry of Education, Ethiopia [2010], Map capacity development design study, Technical report, FDRE.

Ministry of Education, Ethiopia [2015], Education statistics annual abstract 2006 e.c., Technical report, Federal Democratic Republic of Ethiopia (http://www.moe.gov.et/English/Information/Pages/AnnualAbstract.aspx).

Ministry of Finance and Economic Development, Ethiopia [2002], 'Ethiopia: Sustainable development and poverty reduction program', *IMF Policy Paper, http://www.imf.org/external/np/prsp/2002/eth/01/073102.pdf,* .

Ministry of Finance and Economic Development, Ethiopia [2010], Growth and transformation plan 2010/11 - 2014/15, Technical report, Federal Democratic Republic of Ethiopia.

Mookherjee, D. [2006], 'Decentralization, hierarchies, and incentives: A mechanism design perspective', *Journal of Economic Literature* **44**(2), 367–390.

Moulin, H. [1984], 'Dominance solvability and cournot stability', *Mathematical Social Sciences* **7**(1), 83–102.

Muralidharan, K., Niehaus, P. and Sukhtankar, S. [2016], 'Building state capacity: Evidence from biometric smartcards in india', *American Economic Review* **106**(10), 2895–2929.

Muralidharan, K. and Sundararaman, V. [2011], 'Teacher performance pay: Experimental evidence from india', *Journal of Political Economy* **119**(1), 39–77.

Myatt, D. P. and Wallace, C. [2011], 'Endogenous information acquisition in coordination games', *The Review of Economic Studies* **79**(1), 340–374.

Myerson, R. [1981], 'Optimal auction design', *Mathematics of Operations Research* **6**(1), 58–73.

Nath, A. [2015], 'Bureaucrats and politicians: How does electoral competition affect bureaucratic performance?', *Institute for Economic Development (IED) Working Paper* **269**.

Oates, W. E. [1999], 'An essay on fiscal federalism', *Journal of Economic Literature* **37**(3), 1120–1149.

OATES, W. E. [2005], 'Toward a second-generation theory of fiscal federalism', *International Tax and Public Finance* **12**, 349–373.

OECD [2017], *Government at a glance 2017*, Organization for Economic Cooperation and Development.

Olken, B. A. [2007], 'Monitoring corruption: Evidence from a field experiment in indonesia', *Journal of Political Economy* **115**(2).

Olken, B. A., Onishi, J. and Wong, S. [2014], 'Should aid reward performance? evidence from a field experiment on health and education in indonesia', *American Economic Journal: Applied Economics* **6**(4), 1–34.

Oppedisano, V. [2014], 'Higher education expansion and unskilled labour market outcomes', *Economics of Education Review* **40**, 205–220.

Oswald, F. and Schenker, P. [2010], 'Nestown: New ethiopian sustainable town a real life experiment', *INSIDE THIS ISSUE* **7**(1/2), 31.

Oyer, P., Schaefer, S. et al. [2011], 'Personnel economics: Hiring and incentives', *Handbook of Labor Economics* **4**, 1769–1823.

Pepinsky, T. B., Pierskalla, J. H. and Sacks, A. [2017], 'Bureaucracy and service delivery', *Annual Review of Political Science* **20**, 249–268.

Perry, J. L. [1996], 'Measuring public service motivation: An assessment of construct reliability and validity', *Journal of Public Administration Research and Theory* **6**(1), 5–22.

Petersen, M. A. [2004], 'Information: Hard and soft'.

Pomeranz, D. [2015], 'No taxation without information: Deterrence and self-enforcement in the value added tax', *American Economic Review* **105**(8), 2539–69.

Prendergast, C. [2002], 'Uncertainty and incentives', *Journal of Labor Economics* **20**(S2), S115–S137.

Prendergast, C. and Topel, R. [1993], 'Discretion and bias in performance evaluation', *European Economic Review* **37**(2-3), 355–365.

Prendergast, C. and Topel, R. H. [1996], 'Favoritism in organizations', *Journal of Political Economy* **104**(5), 958–978.

Qian, J., Strahan, P. E. and Yang, Z. [2015], 'The impact of incentives and communication costs on information production and use: Evidence from bank lending', *The Journal of Finance* **70**(4), 1457–1493.

Quadrini, V. and Trigari, A. [2007], 'Public employment and the business cycle', *The Scandinavian Journal of Economics* **109**(4), 723–742.

Rasul, I. and Rogger, D. [2013], 'Management of bureaucrats and public service delivery: Evidence from the nigerian civil service', *The Economic Journal* .

Rasul, I. and Rogger, D. [2016], 'Management of bureaucrats and public service delivery: Evidence from the nigerian civil service', *The Economic Journal* .

Rasul, I., Rogger, D. and Williams, M. J. [2018*a*], 'Management and bureaucratic effectiveness: Evidence from the ghanaian civil service'.

Rasul, I., Rogger, D. and Williams, M. J. [2018*b*], 'Management and bureaucratic effectiveness: Evidence from the ghanaian civil service', *Working Paper* .

Ravishankar, V., Kello, A. and Tiruneh, A. [2010], 'Ethiopia: Education public expenditure review', *Ministry of Education, Addis Ababa* .

Rogger, D. [2014], 'The causes and consequences of political interference in bureaucratic decision making: Evidence from nigeria', *Working Paper* .

Rogger, D. [2017], 'Who serves the poor?'.

Rogger, D. and Somani, R. [2018], 'Hierarchy and information'.

Rose-Ackerman, S. [1986], 'Reforming public bureaucracy through economic incentives?', *Journal of Law, Economics, & Organization* **2**(1), 131–161.

Royston, P. [1992], 'Comment on sg3. 4 and an improved d'agostino test', *Stata Technical Bulletin* **1**(3).

Samuelson, P. A. [1954], 'The pure theory of public expenditure', *The Review of Economics and Statistics* pp. 387–389.

Somani, R. [2017], 'Efficiency losses from subjective performance evaluations and supervisor bias: Evidence from the ethiopian civil service', *Working Paper* .
URL: *http://www.rsomani.com/research.html*

Somani, R. [2018], 'Does providing information to civil servants improve service delivery?', *Working Paper* .

Stein, J. C. [2002], 'Information production and capital allocation: Decentralized versus hierarchical firms', *Journal of Finance* pp. 1891–1921.

Stiglitz, J. E. [2017], The revolution of information economics: The past and the future, Technical report, National Bureau of Economic Research.

Sutton, J. and Kellow, N. [2014], 'An enterprise map of ethiopia, international growth centre'.

Tereda, A. [2014], 'Design and implementation of balanced scorecard in the development bank of ethiopia'.

Todaro, M. P. [1969], 'A model of labor migration and urban unemployment in less developed countries', *The American Economic Review* **59**(1), 138–148.
URL: *http://www.jstor.org/stable/1811100*

Uchida, H., Udell, G. F. and Yamori, N. [2012], 'Loan officers and relationship lending to smes', *Journal of Financial Intermediation* **21**(1), 97–122.

UNESCO [2015], Education for all 2015 national review report: Ethiopia, Technical report, UNESCO.

United Nations [2015], 'Millennium development goals'.
URL: *http://www.un.org/millenniumgoals/*

United Nations Educational, Scientific and Cultural Organization [2012], Ethiopia efa profile, Technical report, UNESCO.

Villanueva, C. [2003], Education management system (emis) and the formulation of education for all (efa) plan of action, 2002-2015, Technical report, UNESCO.

Wooldridge, J. M. [2010], *Econometric analysis of cross section and panel data*, MIT press.

World Bank [2010*a*], 'Ethiopia - public finance review 2010', *Public Expenditure Review World Bank* .

World Bank [2010*b*], Woreda and city administrations benchmarking survey iii, Technical report, World Bank.

World Bank [2013*a*], Ethiopia general education quality improvement project ii, Technical report, World Bank.

World Bank [2013*b*], 'World bank supports ethiopia's plan to transform education for more than 21 million children', Online.
URL: *http://www.worldbank.org/en/news/press-release/2013/12/02/world-bank-ethiopia-transform-education-children*

World Bank [2015*a*], Ethiopia poverty assessment 2014, Technical report, World Bank.
URL: *https://openknowledge.worldbank.org/handle/10986/21323*

World Bank [2015*b*], 'Expenditure data', Online.
URL: *http://data.worldbank.org/indicator/SE.XPD.TOTL.GB.ZS*

World Bank [2015*c*], 'General education quality improvement program', Online.
URL: *http://www.worldbank.org/projects/P129828/ethiopia-general-education-quality-improvement-project-ii?lang=en*

World Bank [2015*d*], 'Public finance review ethiopia', *Forthcoming* .

World Bank [2016*a*], 'Ethiopia public expenditure review', *World Bank Group: Washington, DC* .

World Bank [2016*b*], 'World development report 2016: Digital dividends', *World Bank Group: Washington, DC* .

World Bank [2016*c*], 'World development report 2016: Digital dividends', *World Bank Group: Washington, DC* .

World Bank [Forthcoming], 'Moving further on civil service reforms in ethiopia', *World Bank Working Paper Series* .

Zike, W. M. and Ayele, B. K. [2015], 'State of service delivery in ethiopian primary schools: Findings from the ethiopia education service delivery indicator survey', *Middle Eastern & African Journal of Educational Research MAJER Issue: 16* p. 35.